Stanford University Library
JAN
16
1902

JANUARY 1902

# ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

ORMOND STONE W. E. BYERLY H. S. WHITE
W. F. OSGOOD F. S. WOODS

PUBLISHED UNDER THE AUSPICES OF HARVARD UNIVERSITY

SECOND SERIES Vol. 3 No. 2

FOR SALE BY
THE PUBLICATION OFFICE OF HARVARD UNIVERSITY
2 University Hall, Cambridge, Mass., U. S. A.

London: LONGMANS, GREEN & Co.
39 Paternoster Row

Leipzig: OTTO HARRASSOWITZ
Querstrasse 14

# SOME APPLICATIONS OF THE METHOD OF ABRIDGED NOTATION.

By Maxime Bôcher.

**1. Introduction.** In proving a geometrical proposition by the method of abridged notation certain identities, often of an extremely simple character, are used, from whose geometric interpretation the proof of the theorem follows. It frequently happens that the set of identities we have used in the proof of our theorem admits of a second geometrical interpretation, and when we have noticed this fact, we have before us a new geometrical theorem which we have, so to speak, deduced from the theorem from which we started, or rather from our method of proving it.

To illustrate what I mean by a very simple example, let us consider the theorem that the bisectors of the angles of a triangle meet in a point. To prove this by the method of abridged notation we indicate the equations of the sides of the triangle by

$$u = 0, \qquad v = 0, \qquad w = 0. \tag{1}$$

If we suppose that all these equations are written in the normal form:

$$x \cos \alpha + y \sin \alpha - p = 0,$$

the equations of the bisectors of the angles will be

$$u - v = 0, \qquad v - w = 0, \qquad w - u = 0; \tag{2}$$

and now the identity:

$$(u - v) + (v - w) + (w - u) \equiv 0 \tag{3}$$

shows us at once that these lines meet in a point.

Suppose, however, instead of interpreting the equations (1) as the equations of three straight lines, we regard them as the equations of three circles all written in the normal form:

$$x^2 + y^2 + ax + by + c = 0.$$

Then the equations (2) will represent the common chords of these circles, and the identity (3) tells us that these common chords meet in a point.

Thus we see that the method of abridged notation may be used to deduce new theorems from known ones, just as the methods of projection, reciprocation, and inversion are commonly used. As this fact has, in spite of its elementary nature, gained as yet no place, so far as I have found, in even the best text books,* I propose to illustrate it here by first reproducing the well known proofs by the method of abridged notation of the harmonic properties of the complete quadrilateral and of Desargues's theorem, and then interpreting these proofs in such a way as to obtain theorems which, instead of referring to rectilinear figures, as did the original propositions, refer to much more general curvilinear figures. In doing this I have paid special attention to the case in which these curved lines are circles, not only because this is the simplest case, but because the theorems thus obtained are suited to hold a fundamental position, as I expect to show on another occasion, in a systematic development of what has been termed the Geometry of Inversion.† It must, however, be clearly understood that the generalizing process which we use in this simple case is not identical with the process of inversion, but that the results obtained include as a very special case the theorems which the method of inversion would give us.

**2. Rectilinear Figures.** Let us denote the sides of a complete quadrilateral by the equations:

$$\alpha = 0, \quad \beta = 0, \quad \gamma = 0, \quad \delta = 0. \tag{4}$$

Denoting the six vertices by

$$(\alpha, \beta), \quad (\gamma, \delta), \quad (\alpha, \gamma), \quad (\beta, \delta), \quad (\alpha, \delta), \quad (\beta, \gamma), \tag{5}$$

we will let

$$\begin{aligned} u = 0 &\text{ represent the diagonal joining } (\beta, \gamma) \text{ and } (\alpha, \delta), \\ v = 0 &\quad " \quad " \quad " \quad " \quad (\alpha, \gamma) \quad " \quad (\beta, \delta), \\ w = 0 &\quad " \quad " \quad " \quad " \quad (\alpha, \beta) \quad " \quad (\gamma, \delta). \end{aligned} \tag{6}$$

* Except in the case where we merely pass from point to line coordinates,—a case which is essentially equivalent to reciprocation.

† *Cf.* Klein's *Erlanger Programm* (1872), reprinted in English translation and with a few additions in the *Bulletin of the New York Mathematical Society* for July 1893. A German reprint with still further additions will be found in the *Math. Annalen*, vol 43 (1893), p. 63.

The values of $u$, $v$, $w$, in terms of $\alpha$, $\beta$, $\gamma$, $\delta$ may be obtained as follows:— We may write the equation of the diagonal $u$ in either of the forms

$$b\beta + c\gamma = 0, \qquad -a\alpha - d\delta = 0.$$

If we so choose the constants $a$, $b$, $c$, $d$ that the first members of these equations are equal, not merely proportional, we get the identity:

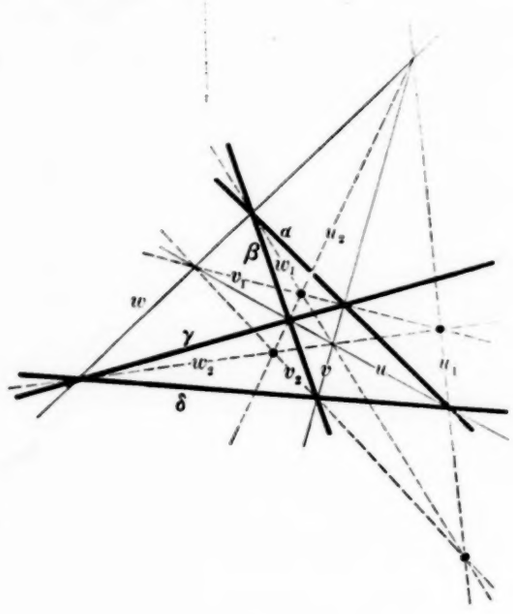



FIG. 1

$$(7_1) \qquad u \equiv b\beta + c\gamma \equiv -a\alpha - d\delta.$$

From this it follows that

$$b\beta + d\delta \equiv -a\alpha - c\gamma.$$

If we set this expression equal to zero we get, as we see from the left hand side, a line through $(\beta, \delta)$, as we see from the right hand side, a line through $(\alpha, \gamma)$. We may therefore write:

$$(7_2) \qquad v \equiv b\beta + d\delta \equiv -a\alpha - c\gamma.$$

Similarly we get:

$$(7_3) \qquad w \equiv c\gamma + d\delta \equiv -a\alpha - b\beta.$$

Finally indicating the lines which join the vertices of the quadrilateral with the points of intersection of the diagonals, as shown in the figure, by:

$$(8)\quad u_1 = 0, \qquad u_2 = 0, \qquad v_1 = 0, \qquad v_2 = 0, \qquad w_1 = 0, \qquad w_2 = 0,$$

we have

$$(9)\quad \begin{aligned} u_1 &\equiv v + w \equiv -a\alpha + d\delta, \\ u_2 &\equiv v - w \equiv \quad b\beta - c\gamma, \\ v_1 &\equiv w + u \equiv -a\alpha + c\gamma, \\ v_2 &\equiv w - u \equiv \quad d\delta - b\beta, \\ w_1 &\equiv u + v \equiv -a\alpha + b\beta, \\ w_2 &\equiv u - v \equiv \quad c\gamma - d\delta. \end{aligned}$$

From these formulæ we read off at once the harmonic properties of the complete quadrilateral, for instance that the two sides $\alpha$ and $\beta$ are separated harmonically by the diagonal $w$ through their point of intersection and the line $w_1$ which joins their point of intersection with the intersection of the two other diagonals (*cf.* ($7_3$) and the fifth formula (9) ).

We may add that formulæ (9) also show that the six lines (8) pass three by three through four points, thus forming the sides of a complete quadrangle.

We pass now to Desargues's theorem:—

*If two triangles are so situated that the points of intersection of corresponding sides are collinear, then the lines joining corresponding vertices are concurrent.*

To prove this theorem we denote the sides of the first triangle by:

$$(10)\qquad u_1 = 0, \qquad u_2 = 0, \qquad u_3 = 0;$$

the sides of the second by:

$$(11)\qquad v_1 = 0, \qquad v_2 = 0, \qquad v_3 = 0.$$

The line on which the points of intersection of corresponding sides lie may be represented by an equation in any one of the following forms:

$$a_1 u_1 + b_1 v_1 = 0, \qquad a_2 u_2 + b_2 v_2 = 0, \qquad a_3 u_3 + b_3 v_3 = 0.$$

Accordingly, by properly choosing the constants, we can make the first members of these equations identically equal:

$$(12)\qquad a_1 u_1 + b_1 v_1 \equiv a_2 u_2 + b_2 v_2 \equiv a_3 u_3 + b_3 v_3.$$

From this we get at once the following three identities:

$$
\begin{aligned}
w_3 &\equiv a_1u_1 - a_2u_2 \equiv -b_1v_1 + b_2v_2,\\
w_1 &\equiv a_2u_2 - a_3u_3 \equiv -b_2v_2 + b_3v_3,\\
w_2 &\equiv a_3u_3 - a_1u_1 \equiv -b_3v_3 + b_1v_1.
\end{aligned} \tag{13}
$$

From these identities we infer that the equations:

$$w_1 = 0, \qquad w_2 = 0, \qquad w_3 = 0 \tag{14}$$

represent the lines joining corresponding vertices of the two triangles, and since

$$w_1 + w_2 + w_3 \equiv 0, \tag{15}$$

these lines are concurrent.

**3. Generalization to Circles.** Let us now try to interpret the analytical work of the preceding section on the supposition that the equations there used represent not straight lines, but circles.

We begin with the theorems about the complete quadrilateral.

The figure formed by four circles, no three of which meet in the same two points, we will call a complete circular quadrilateral, the circles (4) being the *sides* of the quadrilateral and the six pairs of points (5) being its six *sets of vertices*. These latter arrange themselves in pairs of *opposite* sets of vertices. The equations $b\beta + c\gamma = 0$ and $-a\alpha - d\delta = 0$ now represent circles through the pairs of points $(\beta, \gamma)$ and $(\alpha, \delta)$ respectively, but it will not be possible in general to make these two circles coincide, since a circle cannot in general be passed through two pairs of points.

In order that our analytic work be applicable to this case it is therefore necessary that the four points $(\beta, \gamma)$ and $(\alpha, \delta)$ be concyclic (*i. e.* lie on a circle). This circle will then be represented by the equation $u = 0$ (*cf.* $(7_1)$). From this equation we infer as before that $v = 0$ and $w = 0$ (*cf.* $(7_2)$ and $(7_3)$) represent circles through the two pairs of points $(\alpha, \gamma)$, $(\beta, \delta)$ and $(\alpha, \beta)$, $(\gamma, \delta)$ respectively. We thus get the theorem:

*If in a complete circular quadrilateral two opposite sets of vertices are concyclic, the same will be true of each of the other pairs of opposite sets of vertices.**

If a circle passes through a pair of opposite sets of vertices of a complete

* Another form of statement for this theorem which brings it into relation with Desargues's theorem is the following, in which we denote the figure formed by two circles as a *crescent*, the two circles being the *sides* and their points of intersection the *vertices*:

*If two crescents are so situated that their vertices are concyclic, the points of intersection of*

circular quadrilateral, we will call it a diagonal of this quadrilateral. We may then state the last theorem as follows:

*If a complete circular quadrilateral has one diagonal, it has three.*

We will recall here two terms which we shall find it convenient to use. $S_1 = 0$ and $S_2 = 0$ being any two given circles, we call the simply infinite family

$$k_1 S_1 + k_2 S_2 = 0 \tag{16}$$

a *pencil* of circles: and we use the same term for a finite number of circles of this family. Similarly $S_1 = 0$, $S_2 = 0$, $S_3 = 0$ being any three circles which do not belong to a pencil the doubly infinite family:

$$k_1 S_1 + k_2 S_2 + k_3 S_3 = 0 \tag{17}$$

we call a *net*.*

We now infer at once from formula (7):

*A necessary and sufficient condition that a complete circular quadrilateral have diagonals is that its sides belong to a net. If this condition is fulfilled, the diagonals belong to the same net.*

Under these conditions any two sides of the quadrilateral (as $\alpha, \beta$) and the two diagonals which do not pass through the intersections of these sides (as $u, v$) form a complete circular quadrilateral whose sides belong to a net and which therefore has diagonals. Two of these diagonals are the other two sides $(\gamma, \delta)$ of the original quadrilateral, but the third is the circle $(w_1)$ through the points of intersection of $\alpha$ with $\beta$ and of $u$ with $v$. Similarly we obtain the other circles (8).

By the *cross ratio* of a pencil of four circles we shall understand the cross ratio of the pencil of tangents at any one of their points of intersection.† With

---

*corresponding sides (the correspondence being taken in either one of the two possible ways) are also concyclic.*

Still another form, in which, however, only half the theorem is explicitly stated, is the following:

*Through each of the sets of vertices of a circular triangle a circle is passed. If these three circles meet in a first point, they will meet again in a second point.*

As thus stated the theorem is seen to be identical with the generalization by inversion of the theorem that the common chords of three circles meet in a point.

* We mention in passing that such a net consists of all circles which cut at right angles the common orthogonal circle of $S_1$, $S_2$, $S_3$. *Cf.* for instance Casey's *Analytical Geometry*, p. 107.

† Or more generally the cross ratio of the pencil formed by the polars of an arbitrarily chosen point. This more general definition is necessary if the circles of the pencil touch each other.

this definition it is readily seen that a necessary and sufficient condition that the circles $S_1$ and $S_2$ be divided harmonically by $S_1 + \lambda S_2$ and $S_1 + \mu S_2$ is that $\lambda = -\mu$. We thus obtain, precisely as in §1, the harmonic properties of the complete circular quadrilateral, of which the following is typical:

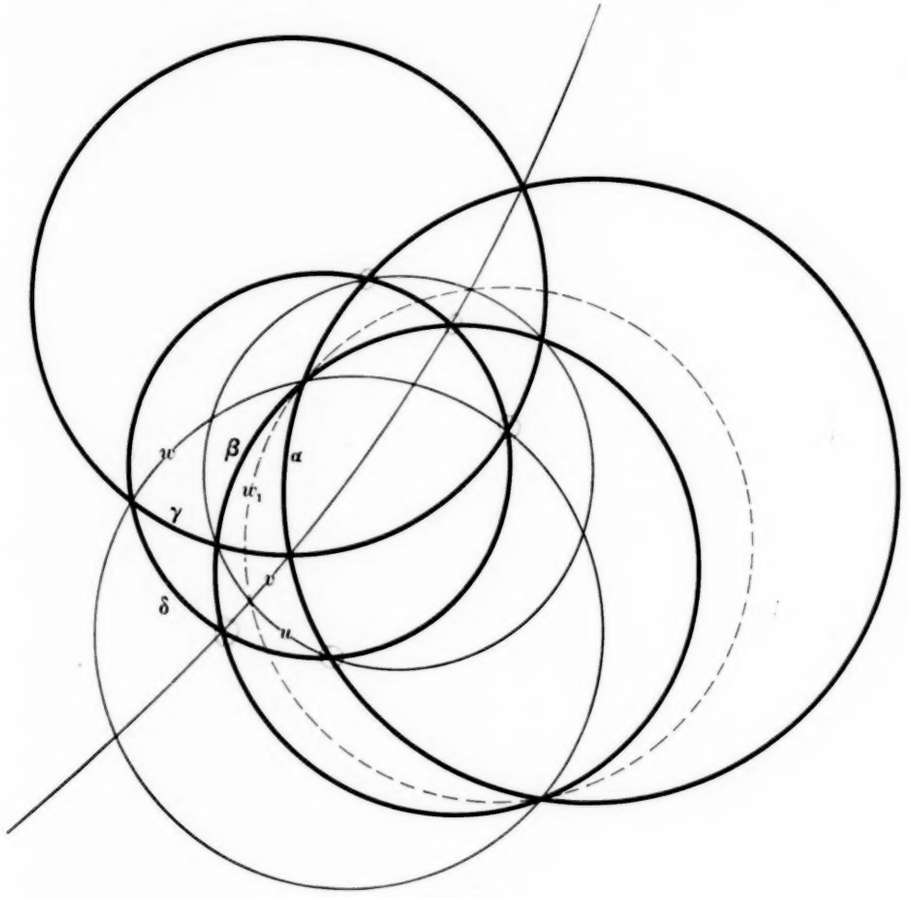



FIG. 2.

*In a complete circular quadrilateral whose sides belong to a net, any two sides are divided harmonically by the diagonal through their points of intersection, and the circle through their points of intersection and through the points of intersection of the other two diagonals.*

We leave to the reader the extension of the theorem that the lines (8) form a complete quadrangle.

Up to this point all the circles we have used have belonged to a single net;* and it might seem at first sight as though the method we are using must necessarily always give us theorems which involve only circles of a single net, since the straight lines from which these circles were obtained themselves form a net. This, however, is not the case, as the extension of Desargues's theorem will now show. For although any four lines of the plane are connected by a relation of the form $a\alpha + b\beta + c\gamma + d\delta \equiv 0$, it is not at all necessary that this relation be involved in the formulæ we use in proving our proposition, as is shown by the proof we have given of Desargues's theorem.†

Turning now to the proof of Desargues's theorem given in §1 we have to consider two circular triangles (10) and (11). We understand of course by a circular triangle the figure formed by three circles which do not form a pencil. Such a triangle has three sides and three pairs of vertices. The analysis of §1 then gives us the theorem:

*If two circular triangles are so situated that the three pairs of points of intersection of corresponding sides are concyclic, then any pair of vertices of the first triangle and the corresponding pair of vertices of the second are concyclic,‡ and the three circles thus determined form a pencil.*

The ten circles involved in this theorem do not necessarily all belong to a net, since the three sides of the first triangle and one side of the second may be arbitrarily chosen.

**4. Curves of the n-th Order.** We will now regard the equations from which we started as representing plane curves of the $n$th order instead of straight lines as in §2 or circles as in §3. We will speak of pencils and nets

* In fact our formulæ practically establish a one-to-one transformation of the straight lines of the plane into the circles of a net, and therefore a one-to-two point transformation (a point determined by two straight lines into the pair of points determined by the corresponding circles). Such a correspondence can easily be established geometrically by projecting the plane from an arbitrarily chosen point onto an arbitrarily chosen spherical surface, and then projecting back stereographically onto the plane.

† A different proof of this theorem which does not have this advantage is given in Salmon's *Conic Sections*, p. 59.

‡ This first part of the theorem does not go beyond what we have already proved, since two pairs of corresponding sides of the triangle form a complete circular quadrilateral lying wholly in a net.

of curves of the $n$th order, using the words in the ordinary sense similar to that explained for circles in §3. We will speak of four curves of the $n$th order, no three of which form a pencil, as the *sides* of a *complete curvilinear quadrilateral.** Each pair of sides of this quadrilateral determines a pencil, and if the pencils determined by two opposite pairs of sides have a curve in common, we will speak of this curve as a *diagonal* of the quadrilateral. Formulæ (7) now establish the following theorem:

*A necessary and sufficient condition that a complete curvilinear quadrilateral have a diagonal is that its sides belong to a net, and if this condition is fulfilled, it has three and only three diagonals which also belong to the net.*

Let us now consider the pencil of four curves

$$S_1 = 0, \qquad S_2 = 0, \qquad S_1 + \lambda S_2 = 0, \qquad S_1 + \mu S_2 = 0.$$

The linear polars of any point $P$ of the plane, excluding possibly certain exceptional positions for $P$, are easily shown to form a pencil, and the cross ratio of this pencil is seen to be independent of the position of $P$ and to have the value $\lambda/\mu$. This quantity we will speak of as the cross ratio of the pencil of curves.† In particular we may in general‡ take for $P$ a point of intersection of the curves, in which case the cross ratio of the pencil of curves is given by the cross ratio of the pencil of tangents at this point of intersection.

If $\lambda = -\mu$ we say that the curves $S_1$ and $S_2$ are *divided harmonically* by the other two curves.

Formulæ (7) and (9) now give us a number of theorems concerning the complete curvilinear quadrilateral, of which we state the following one:

*If a complete curvilinear quadrilateral is formed by four curves of a net, any two sides $\alpha$, $\beta$ of the quadrilateral are divided harmonically by the diagonal belonging to the pencil $(\alpha, \beta)$ and the curve common to the pencil $(\alpha, \beta)$ and the pencil determined by the other two diagonals.*

The extension to this case of the further theorems contained in §2 we leave to the reader.

---

* Such a quadrilateral has in general six sets of $n^2$ vertices each. In case some of the sides are reducible curves two sides may, however, have a whole curve in common. For this reason we have avoided speaking of vertices in the text.

† If instead of considering the linear polars of $P$ we consider the polars of any order, it is clear that these will also form a pencil having the same cross ratio as that of the original pencil.

‡ Provided, namely, there exists a point of intersection where neither $S_1$ nor $S_2$ has a double point, and where these curves do not touch each other.

**5. Extension to Surfaces.** It is hardly necessary to do more than mention the possibility of interpreting the analysis of §2 on the supposition that the equations there used represent surfaces in space instead of lines or curves in the plane. Thus, for instance, if we regard the equations as representing spheres, the formulæ concerning the complete quadrilateral in §2 give us certain theorems concerning spheres which have two points in common, while the proof of Desargues's theorem gives us a theorem concerning ten spheres which cut a given sphere orthogonally. The statement of these theorems and of similar theorems concerning surfaces of any order presents no difficulty.

Even without changing the degrees of the equations we use, the method here considered gives us the means of transferring theorems in plane geometry to the case of three dimensions. The theorems of §2 are too simple to give anything interesting in this way; but it will be a good exercise for a student to transfer by this method to quadric surfaces the theorem that *if three conics have double contact with a fourth, six of their chords of intersection pass three by three through four points*, using for this purpose the proof by abridged notation given in Salmon's *Conic Sections*, p. 243.

It will be seen that examples of cases to which our method applies can be multiplied indefinitely. We have confined ourselves here to a few of the very simplest illustrations.

HARVARD UNIVERSITY,
NOVEMBER 1901.

# ON THE ROOTS OF FUNCTIONS CONNECTED BY A LINEAR RECURRENT RELATION OF THE SECOND ORDER.*

By M. B. Porter.

At the close of his first memoir on homogeneous linear differential equations of the second order in the first volume of Liouville's Journal, Sturm remarks that results analogous to those obtained for the differential equation were first discovered by him while considering solutions of a homogeneous equation in finite differences of the form,

$$L_n y_{n+1} + M_n y_n + N_n y_{n-1} = 0, \tag{1}$$

when $L$, $M$, and $N$ are functions, subject to certain restrictions, of the positive integral index $n$ and a variable parameter, and that it was while prosecuting this inquiry that he discovered his famous theorem concerning the isolation of the roots of a polynomial. By passing from the discrete index $n$ to the continuous variable $x$, Sturm first arrived at the theorems for the differential equation in the paper referred to. This earlier method of dealing with the problems solved by Sturm was perhaps regarded by him as merely heuristic, or at least as less elegant than the methods finally employed, and his earlier researches were never published.

Characterizing a sequence of real functions $y_0, y_1, \ldots y_n$ as *Sturmian*† when the difference in the number of variations of sign in the sequence for two particular real values of the argument is equal to the number of real roots of $y_n$ in the interval delimited by them, it is certain that Sturm determined at least one simple sufficient condition to which a sequence must conform in order that it may be *Sturmian*, and that he investigated the inverse problem of constructing such sequences by means of a recurrent relation of type (1). Looked at from the standpoint of the recurrent relation, solutions of (1) enjoying the *Sturmian* property may be called *Sturmian solutions*: undoubtedly Sturm discovered many properties of such solutions.

---

* Read before the American Mathematical Society, February, 1901, under a different title.

† This adjective is here used in a more general sense than in the ensuing sections. *Cf.* the definitions there given.

It is our purpose in the first five sections to reproduce in part these unpublished theorems and in the last section to show how, by means of the Cauchy-Lipschitz theorem for the existence of solutions of a differential equation, it is possible to establish rigorously the analogous theorems, so far as they exist, for the homogeneous linear differential equation of the second order. As has already been stated, in this case the index $n$ is replaced by the independent variable and the variable parameter either figures in the differential equation itself, or in the arbitrary parameters of integration, or in both.

In conclusion it may be said that it is highly probable that Sturm considered a somewhat more general difference equation than (1′) of §3 and so arrived at criteria of a more general character. This point, however, I hope to consider elsewhere.

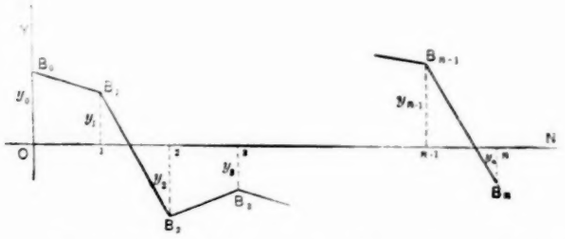
**1. Sturmian Sequences in General.** We begin by considering a series of functions

$$y_n(x), \quad y_{n-1}(x), \ \ldots \ldots \ y_0(x)$$

of the real variable $x$, real, single valued, and analytic in an interval $x_0 \leqq x \leqq x_1$.*

A. *It is further supposed that for no point of this interval is $y_0(x)$ equal to zero.*

It will be convenient to employ the following graph in considering the




above sequence. The ordinate $i\,B_i$ representing $y_i$ in magnitude and sign, the abscissa $Oi$ may conveniently be taken equal to the index $i$. In the graph just described, the intersections of the broken line $B_0\,B_1 \ldots . B_n$ with the $N$-axis, which we shall call *$v$-points*, correspond to the *variations of sign* in the sequence $y_0, y_1, \ldots y_n$.

If now we suppose the parameter $x$ to vary continuously, this broken line will undergo a continuous deformation, the point $B_0$ remaining always above

* Throughout the rest of this paper we shall suppose that the variable $x$ is confined to such an interval.

the origin, while the $v$-points are displaced in a manner which will depend on the particular functions $y_1, \ldots y_n$. It is evident that at least one variation will be lost or gained when $y_n$ changes sign. The vanishing of an intermediate $y$ can cause the loss or gain of a variation only when $B_0 B_1 \ldots . B_n$ is deformed in such a manner that, when a given $y$ vanishes, the adjacent $y$'s have the same sign, or one of the adjacent $y$'s is zero. In order that the number of variations of sign in the sequence $y_0, y_1, \ldots . y_n$ should change as $x$ varies from $x_0$ to $x_1$ *only* when $y_n$ changes sign, it will be *sufficient*, if we impose the condition:

*B. No two consecutive y's can vanish for the same value of x, and when any intermediate y vanishes the adjacent y's have opposite signs.*

A series of functions of the kind specified, *i. e.* fulfilling Conditions *A* and *B*, are said to form a *generalized Sturmian* sequence.*

Let us now consider what further restrictions must be imposed on our sequence in order that it should lose a variation each time $y_n$ vanishes. Let $x'$ be a value of $x$ for which $y_n = 0$. The sequence will lose a variation as $x$ increases through the value $x'$ when and only when $y_n$ and $y_{n-1}$ have opposite signs for values of $x$ a little smaller than $x'$ and the same sign for values a little larger, *i. e.* when

$$(2) \qquad \frac{y_n(x' - \Delta x)}{y_{n-1}(x' - \Delta x)} < 0 < \frac{y_n(x' + \Delta x)}{y_{n-1}(x' + \Delta x)} \qquad \left(\begin{array}{l}\Delta x = \text{a sufficiently small} \\ \text{positive quantity.}\end{array}\right)$$

From (2) we see that $y_n/y_{n-1}$ is increasing with $x$ at the point $x'$, and therefore we have as a consequence of (2):

$$(3) \qquad \left[\frac{d}{dx}\left(\frac{y_n}{y_{n-1}}\right)\right]_{x=x'} \geqq 0.$$

On the other hand, if a variation is to be gained:

$$\left[\frac{d}{dx}\left(\frac{y_n}{y_{n-1}}\right)\right]_{x=x'} \leqq 0.$$

We shall treat explicitly only the case when variations are lost as $x$ increases.

It is to be noted that condition (2) is more general than (3) and that we cannot pass back to (2) except when the inequality sign holds. It is thus a

* *Cf.* Netto. *Algebra*, vol. 1, p. 238.

sufficient, but not a necessary, condition for the loss of a variation that*

$$(4) \qquad \frac{d}{dx}\left(\frac{y_n}{y_{n-1}}\right) = \frac{y'_n}{y_{n-1}} > 0 \qquad \text{when } y_n = 0.$$

When this condition is fulfilled as well as conditions $A$ and $B$, the sequence is called a *Sturmian* sequence.† It should be noted that condition (4) makes it impossible for $y_n$ to have a multiple root in the interval $(x_0, x_1)$ considered.

Starting with a real polynomial which has no multiple roots in $(x_0, x_1)$ and choosing $y_{n-1} \equiv y'_n$, Sturm's modified algorithm of the Greatest Common Divisor leads to a sequence possessing these properties, the ordinary Sturm's Functions of text books on the Theory of Equations.

**2. The Difference Equation.** If $y_n$ denote a function of the integral index $n$, where we shall always suppose $n$ to denote a positive integer or zero, $y_n$ is said to satisfy a *homogeneous linear recurrent relation of the second order* if we have for all values of $n$,

$$(1) \qquad L_n y_{n+1} + M_n y_n + N_n y_{n-1} = 0,$$

where $L_n$, $M_n$, and $N_n$ denote functions of the index $n$. We will assume in what follows that none of the functions $L_n$ is zero.

If we write

$$\begin{aligned} y_{n+1} - y_n &= \Delta y_n, \\ y_{n+2} - y_{n+1} &= \Delta y_{n+1}, \\ \text{and} \qquad \Delta y_{n+1} - \Delta y_n &= \Delta^2 y_n, \end{aligned}$$

relation (1) may be written in the form of a difference equation

$$L'_n \Delta^2 y_n + M'_n \Delta y_n + N'_n y_n = 0,$$

a form entirely analogous to that of the homogeneous linear differential equation of the second order. We shall however employ this equation only in the recurrent form.

If to $y_0$ and $y_1$ be assigned any values we choose,‡ all subsequent $y$'s are uniquely determined by means of (1), since we have assumed that $L_n$ does not vanish; so that (1) may be said to have $\infty^2$ solutions.

If $y_n^{(1)}$ and $y_n^{(2)}$ denote any two solutions, then, $C_1 y_n^{(1)} + C_2 y_n^{(2)}$ is also a

* If multiple roots are excluded (4) is of course as general as (2).

† *Cf.* Weber, *Algebra*, vol. 1, p. 272 (1st ed.).

‡ This is analogous to the determination of a solution of a differential equation of the second order by its value and the value of its derivative at a given point.

solution.* $y_n^{(1)}$ and $y_n^{(2)}$ are said to be linearly dependent if a relation of the form

$$C_1 y_i^{(1)} + C_2 y_i^{(2)} = 0, \qquad i = 0, 1, 2 \ldots n$$

exist between them, where the $C$'s denote constants.

We will now establish the theorem:

*The necessary and sufficient condition for the linear dependence of two solutions $y_n^{(1)}$ and $y^{(2)}$ is*

$$y_0^{(1)} y_1^{(2)} - y_0^{(2)} y_1^{(1)} = 0. \tag{5}$$

For if $y_n^{(1)}$ and $y_n^{(2)}$ are linearly dependent, we have:

$$\begin{aligned} C_1 y_0^{(1)} + C_2 y_0^{(2)} &= 0, \\ C_1 y_1^{(1)} + C_2 y_1^{(2)} &= 0, \end{aligned} \tag{6}$$

where $C_1$ and $C_2$ are not both zero, and therefore (5) is fulfilled.

Conversely, if (5) is fulfilled, there exist two constants $C_1$ and $C_2$ not both zero which satisfy (6). Now if we have:

$$\begin{aligned} C_1 y_k^{(1)} + C_2 y_k^{(2)} &= 0, \\ C_1 y_{k-1}^{(1)} + C_2 y_{k-1}^{(2)} &= 0, \end{aligned}$$

then by equation (1) we have at once,

$$C_1 y_{k+1}^{(1)} + C_2 y_{k+1}^{(2)} = 0.$$

Therefore starting from (6) we see that:

$$C_1 y_n^{(1)} + C_2 y_n^{(2)} = 0 \qquad (n = 2, 3, 4, \ldots).$$

If then $y_n^{(1)}$ and $y_n^{(2)}$ denote any two linearly independent particular solutions and $y_n$ the general solution, and if we determine $C_1$ and $C_2$ by means of the equations,

$$\begin{aligned} y_0 &= C_1 y_0^{(1)} + C_2 y_0^{(2)}, \\ y_1 &= C_1 y_1^{(1)} + C_2 y_1^{(2)}, \end{aligned}$$

then

$$y_n = C_1 y_n^{(1)} + C_2 y_n^{(2)} \qquad (n = 0, 1, 2, 3, \ldots)$$

*i. e. any solution can be expressed linearly in terms of any two particular linearly independent solutions.*

In the next section we shall be mainly concerned with a recurrent relation of the form,

$$y_{n+1} + G_n y_n + y_{n-1} = 0. \tag{1'}$$

* Analogous to the two arbitrary constants in the general solution of a differential equation of the 2nd order.

Equation (1) may be reduced to this form by the change of dependent variable

$$y_n = e_n z_n. \tag{7}$$

Thus (1) becomes

$$L_n e_{n+1} z_{n+1} + M_n e_n z_n + N_n e_{n-1} z_{n-1} = 0.$$

If now we determine $e_n$ by the recurrent relation

$$L_n e_{n+1} = N_n e_{n-1}$$

we get an equation of the desired form, where $e_0$ and $e_1$ may be chosen at pleasure, subject merely to the restriction that they do not vanish, and

$$e_{2n} = \frac{N_{2n-1} N_{2n-3} \cdot \cdot \cdot \cdot N_1}{L_{2n-1} L_{2n-3} \cdot \cdot \cdot \cdot L_1} e_0 \cdot \qquad e_{2n+1} = \frac{N_{2n} N_{2n-2} \cdot \cdot \cdot \cdot N_2}{L_{2n} L_{2n-2} \cdot \cdot \cdot \cdot L_2} e_1,$$

$$G_n = \frac{M_n}{L_n} \frac{e_n}{e_{n+1}} \cdot$$

**3. Sturmian Sequences satisfying a Recurrent Relation of Type (1).** We shall now consider a sequence of functions $y_n (n = 0, 1, \ldots n)$ satisfying a relation of the form,

$$L_n y_{n+1} + M_n y_n + N_n y_{n-1} = 0.$$

in which $L$, $M$, and $N$ are functions of the index $n$ and are real, single valued, and analytic functions of a real variable $x$ in an interval $x_0 \leqq x \leqq x_1$. If, as we shall assume to be the case, $L_n(x)$ does not vanish in this interval, $y_2$ and all $y$'s of higher index will be determined as real single valued analytic functions of $x$ in this interval as soon as $y_0$ and $y_1$ are assigned as arbitrary real analytic functions of $x$.

Let us now choose $y_0$ so that it will not vanish in the interval $(x_0, x_1)$ and seek to determine what conditions may be imposed on $L$, $M$, $N$, $y_0$, and $y_1$, so that the sequence $y_0, y_1, \ldots y_n$ shall form a Sturmian sequence. To ensure that no two consecutive $y$'s shall vanish for the same value of $x$ and that, when any $y$ vanishes, the adjacent $y$'s shall have opposite signs, it will be sufficient if we suppose that *not only $L$ but also $N$ does not vanish in the interval $(x_0, x_1)$ and that these functions have like signs, i. e.* that $LN$ is always positive. If these conditions are satisfied, the $y$'s form a generalized Sturmian sequence.

To determine a sufficient condition that

$$\left[\frac{d}{dx}\left(\frac{y_n}{y_{n-1}}\right)\right]_{x=x'} > 0,$$

when $x'$ denotes a root of $y_n$ in the interval $(x_0, x_1)$, we shall suppose (1) reduced to the form (1′) by the transformation (7) of §2 and consider the equation as given in the form,

$$y_{n+1} + G_n y_n + y_{n-1} = 0, \tag{1′}$$

when evidently the condition $LN > 0$ is fulfilled.* The relation between $z_n$ and $y_n$ (§2) shows that, under the restrictions placed on $L$ and $N$, $z$ and $y$ vanish together.

Let $x$ take the positive increment $\Delta x$, and let us write $y_i(x + \Delta x) = \bar{y}_i$, etc. Then (1′) becomes

$$\bar{y}_{n+1} + \bar{G}_n \bar{y}_n + \bar{y}_{n-1} = 0.$$

So that, when $y_n \bar{y}_n$ is not zero, we have, since $y_0(x) \neq 0$:

$$\Delta\frac{y_{n+1}}{y_n} = \frac{\bar{y}_{n+1}}{\bar{y}_n} - \frac{y_{n+1}}{y_n} = \frac{1}{\bar{y}_n y_n}\left[(-\bar{G}_n + G_n)\,\bar{y}_n y_n + \bar{y}_n y_{n-1} - y_n \bar{y}_{n-1}\right]$$

$$= \frac{1}{\bar{y}_n y_n}\left[\sum_1^n (-\bar{G}_m + G_m)\,\bar{y}_m y_m + \bar{y}_0 y_0\left(\frac{\bar{y}_1}{\bar{y}_0} - \frac{y_1}{y_0}\right)\right]. \tag{8}$$

Since all the $y$'s are continuous, we can, for a given value of $x$, take $\Delta x$ so small that

$$\bar{y}_m y_m \geqq 0, \qquad m = 0, 1, 2, \ldots n.$$

Let us now impose the restrictions:

$$\left.\begin{array}{ll} 1° & \bar{G}_m \leqq G_m \\ & m = 1, 2 \ldots n. \\ 2° & \bar{y}_1/\bar{y}_0 \geqq y_1/y_0 \end{array}\right\} \qquad x_0 \leqq x \leqq x_1$$

where, if the equality sign holds in 2°, we require either that there be two successive relations 1° where it does not hold, or that it does not hold in the last relation 1°.

* It may be remarked in passing that this transformation is unnecessary if $N/L$ is independent of $x$.

Under these circumstances a quantity $\delta$ can be found such that, if $y_n(x) \neq 0$,

$$\Delta \frac{y_{n+1}}{y_n} > 0 \qquad \text{when } \Delta x < \delta.$$

Thus $y_{n+1}/y_n$ will then increase continually with $x$ from $-\infty$ to $+\infty$ in every interval of $(x_0, x_1)$ delimited by two consecutive roots of $y_n(x)$. Since the $y$'s are analytic functions of $x$, the limits

$$\lim_{\Delta x = 0} \frac{\Delta\left(\frac{y_{n+1}}{y_n}\right)}{\Delta x}, \qquad \lim_{\Delta x = 0} \frac{\Delta\left(\frac{y_1}{y_0}\right)}{\Delta x}$$

exist and we have,

$$\left(\frac{y_{n+1}}{y_n}\right)' = \frac{1}{y_n^2}\left[\sum_1 - G'_m y_m^2 + y_0^2 \left(\frac{y_1}{y_0}\right)'\right] \qquad y_n \neq 0. \tag{9}$$

The conditions

$$\begin{aligned} &1^\circ \quad G'_m(x) \leqq 0, \\ &2^\circ \quad \left(\frac{y_1}{y_0}\right) \geqq 0, \end{aligned} \qquad m = 1, 2, \ldots n$$

the restrictions above stated as to the simultaneous presence of equalities in 1° and 2° holding, yield at once the inequality

$$\left(\frac{y_{n+1}}{y_n}\right)' > C > 0, \qquad y_n \neq 0,$$

when $x$ is supposed to lie in an interval included within an interval delimited by two consecutive roots of $y_n$. Thus, if conditions 1° and 2° are fulfilled, the functions $y_0, y_1, y_2 \ldots y_n$ form a Sturmian sequence. We may call such solutions *Sturmian solutions*.

Let us now consider the case in which $G_n$ and $y_1/y_0$ are real analytic functions of several parameters $x_1, x_2, \ldots$ in the intervals

$$x_0^{(i)} < x_i < x_1^{(i)}, \qquad i = 1, 2, \ldots.$$

We will assume that these functions satisfy the conditions

$$\begin{aligned} &1^\circ \quad \frac{\partial}{\partial x_i} G_m \leqq 0, \\ &2^\circ \quad \frac{\partial}{\partial x_i}\left(\frac{y_1}{y_0}\right) \geqq 0. \end{aligned} \qquad \begin{matrix} m = 1, 2, \ldots n \\ i = 1, 2, \ldots. \end{matrix}$$

Let us first consider the case in which there are only two parameters $x_1$, $x_2$, and

ask ourselves how the roots of $y_{n+1}$, regarded as a function of $x_1$, vary with $x_2$. It is clear that these roots are continuous (in fact analytic) functions of $x_2$ which, since they are never multiple roots, can cease to be real only by disappearing at one end of our interval. The question is whether they increase or decrease with $x_2$. To settle this question we notice that, $x_1$, $x_2$ being a pair of values for which $y_n \neq 0$, the function $y_{n+1}/y_n$ must, according to what we have proved above, increase when $x_1$ or $x_2$ or both increase. If then $x_1'$, $x_2'$ is a pair of values for which $y_{n+1}$ vanishes, and if starting from these values, we wish to change $x_1$, $x_2$ so that $y_{n+1}$ retains the value zero, we must increase one of them and decrease the other; *i. e.* if we regard $y_{n+1}$ as a function of $x_1$, its roots decrease as $x_2$ increases.

Coming back now to the case in which there are more than two parameters, we will assign to all but two of them fixed values. Then by applying the result just obtained we get the theorem:

*If $y_n(x_1, x_2 \ldots)$ be regarded as a function of the variable $x_i$, the roots of $y_n$ will decrease with the increase of the other parameters.*

**4. Theorems concerning Sturmian Solutions.** We have seen that, if $y_n$ denote a Sturmian solution in a certain interval, then $y_n$ has no multiple roots in this interval and that $y_n/y_{n-1}$ increases continually from $-\infty$ to $+\infty$ in any interval delimited by two consecutive roots of $y_{n-1}$. Similarly, it is seen at once that $(y_n + H_n y_{n-1})/y_{n-1}$ will increase continually from $-\infty$ to $+\infty$ in this same interval provided that

$$\frac{d}{dx} H_n \geqq 0.$$

A similar extension can of course be given to the last theorem of §3. From the foregoing we can now deduce the following theorems.

THEOREM 1. *A variation will be lost in the sequence $y_0, y_1, \ldots y_n$ whenever $y_n$ vanishes.*

THEOREM 2. *The roots of $y_n$ and $y_{n-1}$ separate each other*, or more generally *the roots of $y_n + H_n y_{n-1}$ and $y_{n-1}$ separate each other provided that $H_n' \geqq 0$.*

For, since $y_n/y_{n-1}$ increases continually from $-\infty$ to $+\infty$ in any interval delimited by two consecutive roots of $y_{n-1}$, the function $y_n$, which has no multiple roots, will vanish once and only once when $x$ increases through this interval.

THEOREM 3. *If $y_n^{(1)}$ and $y_n^{(2)}$ denote any two linearly independent solutions of (1'), then if $\left(\frac{y_1^{(1)}}{y_0^{(1)}}\right)' > 0$ and $\left(\frac{y_1^{(2)}}{y_0^{(2)}}\right)' > 0$, the roots of $y_n^{(1)}$ and $y_n^{(2)}$ will separate each other.*

From (1') we have

$$y_n^{(1)} y_{n-1}^{(2)} - y_n^{(2)} y_{n-1}^{(1)} \equiv y_1^{(1)} y_0^{(2)} - y_1^{(2)} y_0^{(1)} \equiv C(x)$$

where, by the hypothesis that $y_n^{(1)}$ and $y_n^{(2)}$ are linearly independent, $C$ must be a function of $x$ which does not vanish in the interval $(x_0, x_1)$. Further, since $C(x)$ is continuous, and thus cannot change sign, we may suppose $C(x)$ always positive. Let $\bar{x}_2$ and $x_2$ denote two consecutive roots of $y_n^{(2)}$. By Theorem 2 this interval contains a single root of $y_{n-1}^{(2)}$: hence, since

$$y_n^{(1)}(x_2)\, y_{n-1}^{(2)}(x_2) = C(x_2) > 0$$

and

$$y_n^{(1)}(\bar{x}_2)\, y_{n-1}^{(2)}(\bar{x}_2) = C(\bar{x}_2) > 0,$$

it follows that $y_n^{(1)}$ vanishes at least once in the interval determined by two consecutive roots of $y_n^{(2)}$. Similarly, $y_n^{(2)}$ vanishes at least once in an interval determined by two consecutive roots of $y^{(1)}$, and hence the theorem.

Let $y_n^{(1)}$ and $y_n^{(2)}$ be any two linearly independent solutions, and let us denote by $y_n$ the solution

$$y_n = z y_n^{(1)} + y_n^{(2)}$$

where $z$ is a parameter. Then we have

$$\frac{y_n}{y_{n-1}} = \frac{z y_n^{(1)} + y_n^{(2)}}{z y_{n-1}^{(1)} + y_{n-1}^{(2)}}$$

and

$$\frac{\partial}{\partial z}\left(\frac{y_n}{y_{n-1}}\right) = \frac{y_n^{(1)} y_{n-1}^{(2)} - y_n^{(2)} y_{n-1}^{(1)}}{(z y_{n-1}^{(1)} + y_{n-1}^{(2)})^2} = \frac{y_0^{(1)} y_0^{(2)}}{(z y_{n-1}^{(1)} + y_{n-1}^{(2)})^2}\left(\frac{y_1^{(1)}}{y_0^{(1)}} - \frac{y_1^{(2)}}{y_0^{(2)}}\right).$$

Thus $y_n/y_{n-1}$ increases continually with $z$ between any two consecutive roots of $y_{n-1}$, if $y_1^{(1)}/y_0^{(1)} > y_1^{(2)}/y_0^{(2)}$. If $z$ decrease from a very large value to a very small one, $y_n/y_{n-1}$, starting from a value as near $y_n^{(1)}/y_{n-1}^{(1)}$ as we please, will approach $y_n^{(2)}/y_{n-1}^{(2)}$ as its limit, and the $v$-points of $y_n \ldots y_0$ will, starting from coincidence with those of $y_n^{(1)} \ldots y_0^{(1)}$, all move in the same sense into coincidence with those of $y_n^{(2)} \ldots y_0^{(2)}$. Since $y_n$ always remains linearly independent of $y_n^{(1)}$ and $y_n^{(2)}$, a $v$-point of $y_n \cdots y_0$ could not come into coincidence with one of $y_n^{(2)} \ldots y_0^{(2)}$. If $y_1^{(1)}/y_0^{(1)} < y_1^{(2)}/y_0^{(2)}$ we can reason in the same way, and thus we get

THEOREM 4. *If $y_n^{(1)}$ and $y_n^{(2)}$ denote any two linearly independent solutions of* (4) *such that $y_0^{(1)}$ and $y_0^{(2)}$ do not vanish, no restrictions as to continually increasing or decreasing being imposed on $G_n(x)$, the v-points of the series $y_n^{(1)} \ldots y_0^{(1)}$ and of $y_n^{(2)} \ldots y_0^{(2)}$ will alternate, i. e. the number of variations in the two sequences will differ at most by unity.*

Theorems 3 and 4 are analogous to well known theorems of Sturm for the linear differential equation of the second order. Theorems 1 and 2 seem to possess no analogon for the differential equation.

The analogon of Sturm's *Theorem of Oscillation* may be stated as follows.

THEOREM 5. *If $G_n(x)$ continually increases with $x$ and is such that $G_n(x_0) \geqq L^2$ and $G_n(x_1) \leqq -M^2$ $(n = 0, 1, 2, \ldots)$, and if $L^2$ and $M^2$ are sufficiently large, there is one and only one value of $x (x_0 \leqq x \leqq x_1)$ such that*

1° *The number of variations in the sequence $y_n, y_{n-1}, \ldots y_1$ is equal to an assigned integer $r$ ($r$ not greater than $n-1$) and such that*

2°
$$\frac{y_1(x)}{y_0(x)} = a$$

$$\frac{y_n(x)}{y_{n-1}(x)} = a'$$

*$a$ and $a'$ being any assigned real constants.*

The proof is immediate. We note that, if in the equation

$$\bar{y}_{n+1} + C\bar{y}_n + \bar{y}_{n-1} = 0$$

we take $C$ as a sufficiently large positive constant, the sequence $\bar{y}_n, \bar{y}_{n-1}, \cdots \bar{y}_1$ will present $n-1$ variations of sign and if $C$ is a negative constant of sufficiently large absolute value, this sequence will present no variations of sign. Thus if $L^2$ is sufficiently large, the sequence $y_1, y_2, \ldots y_n$ will present $n-1$ variations of sign. As $x$ increases from $x_0$ towards $x_1$ $G$ will decrease and a variation will be lost every time $y_n$ vanishes: further since the roots of $y_{n-1}$ separate those of $y_n$ and in an interval delimited by two successive roots of $y_{n-1}$, $y_n/y_{n-1}$ increases continually from $-\infty$ to $+\infty$ with the increase of $x$, $y_n/y_{n-1}$ passes in every such interval through any given value, as $\beta$, once and only once. This proves the theorem.

**5. Some Applications.** As examples of sequences which fall at once under the conditions of the preceding sections we will cite first the

*Zonal Harmonics*:

$$P_0(x),\ P_1(x), \ldots . P_n(x): \quad \text{where } P_0 = 1 \text{ and } P_1 = x,$$

while the recurrent relation is

$$(n+1)P_{n+1} = (2n+1)\,xP_n - nP_{n-1}.$$

Since $L_n$ and $N_n$ do not depend on $x$, the conclusions of §3 are valid and we have, the interval $(x_0, x_1)$ extending from $-\infty$ to $+\infty$,

1° The sequence $P_n \ldots P_0$ is *Sturmian*;

2° The roots of $P_n$ and $P_{n-1}$ separate each other;

3° $P_n$ has no multiple roots.

From 1° it follows that all the roots of $P_n$ are real.*

*Lommel's Function*: $R^{n,m}(x)$, defined as follows,

$$J_n(x) = R^{n,m-1}(x)\,J_{n+m-1}(x) - R^{n,m-2}(x)\,J_{n+m}(x),$$

where the $J$'s denote Bessel functions, affords a second example.

Here $R^{n,0}(x) = 1$ and $R^{n,1}(x) = 2(n+1)/x$, the recurrent relation being†

$$R^{n,m+1}(x) = 2\frac{(n+m+1)}{x}R^{n,m}(x) - R^{n,m-1}(x), \tag{10}$$

so that, if $n > -1$, the conditions of §3‡ are fulfilled and $R^{n,0}, R^{n,1}, \ldots R^{n,m}$ form a Sturmian sequence. The roots of $R^{n,m}$ and $R^{n,m-1}$ are all real and alternate and increase with the parameter $n$.

Even if $n < -1$, if $n + m < 0$, the conditions of §3 and the conclusions from them hold.

The conditions at the end of §3 also have place for (10) if $n + m < 0$ and $x$ be changed to $-x$, so that, *if $n$ be increased, the roots of $R^{n,m}(-x) = 0$ decrease; i. e.* the roots of $R^{n,m}(x) = 0$ increase with $n$.

---

* These theorems may be obtained by making use of the relation

$$(x^2-1)P_n' = n\,x\,P_n - n\,P_{n-1}$$

which shows that, if $|x| < 1$, $P_n' P_{n-1} > 0$, when $P_n = 0$ (conditions on p. [57]). *Cf.* Weber's *Algebra*, vol. 1, p. 277.

† *Cf.* Graf and Gubler's *Einleitung in die Theorie der Besselschen Functionen*, vol. 2, p. 102.

‡ The inequalities 1° and 2° of p. 62 are here reversed, but Theorems 1, . . . 5 still subsist if we replace the word *lost* in Theorem 1 by *gained* and reverse the inequalities in Theorem 3.

As a final example we may consider the partial numerators and denominators of the convergents of the continued fraction

$$F(x) = a_1 x + \beta_1 - \cfrac{\gamma}{a_2 x + \beta_2 - \cfrac{\gamma_2}{a_3 x + \beta_3 - }}$$

when the $a$'s are all positive or all negative and the $\gamma$'s are positive functions of the index $n$. Denoting the partial numerators and denominators by $y_n$ and $\bar{y}_n$ respectively we have

$$y_n = (a_n x + \beta_n) y_{n-1} - \gamma_n y_{n-2},$$
$$\bar{y}_n = (a_n x + \beta_n) \bar{y}_{n-1} - \gamma_n \bar{y}_{n-2},$$
$$y_0 = 1, \qquad y_1 = a_1 x + \beta,$$
$$\bar{y}_0 = 0, \qquad \bar{y}_1 = 1, \qquad \bar{y}_2 = a_2 x + \beta_2.$$

Thus the conditions of §3 are fulfilled and $y_n, y_{n-1}, \ldots y_0$ (or $\bar{y}_n, \bar{y}_{n-1}, \ldots \bar{y}_1$) form a Sturmian sequence, all the roots of $y_n(\bar{y}_n)$ are real and are separated by those of $y_{n-1}(\bar{y}_{n-1})$.* From the relation

$$y_n \bar{y}_{n-1} - \bar{y}_n y_{n-1} = \prod_2^n \gamma_n$$

we also infer that the roots of $y_n$ and $\bar{y}_n$ separate each other, *i. e.* the poles and zeros of $y_n/\bar{y}_n$ alternate.

By §3 we conclude that the absolute values of the roots of $y_n(x) = 0$ $(\bar{y}_n(x) = 0)$ decrease with $a_r (r \leqq n)$ if the $a$'s are positive, and increase with the $a$'s if the $a$'s are negative.

The foregoing is the simplest type of continued fraction which satisfies the conditions of §§3 and 4.

The general type suggested by §3 is

$$\phi_0(x) - \cfrac{\gamma_1}{\phi_1(x) - \cfrac{\gamma_2}{\phi_2(x) - \ldots}}$$

where $\phi_n'(x) > 0$.

**6. The Limiting Case of the Differential Equation.** To show that under certain restrictions, to be hereafter given, the difference equation goes over into a differential equation as its limiting form we shall need the following

* This result has been established by Sylvester, *Philosophical Transactions*, vol. 14, part I.

*Lemma.* *If $f(x)$ denote a single valued function of the real variable $x$ possessing finite first and second derivatives in the interval $(x_0, x_1)$ then*•

$$\lim_{\delta=0} \frac{f(x+2\delta) - 2f(x+\delta) + f(x)}{\delta^2} = f''(x), \qquad x_0 < x < x_1.$$

Consider now the difference equation,

$$f_{n+2} + G_n f_{n+1} + f_n = 0 \tag{11}$$

and let

$$f_n = f(a + n\delta), \qquad n\delta = x - a,$$

then (11) may be written

$$\frac{f_{n+2} - 2f_{n+1} + f_n}{\delta^2} = -\left(\frac{2 + G_n}{\delta^2}\right) f_{n+1} \tag{12}$$

If then we suppose that

$$\lim_{\delta=0} \frac{2 + G_n}{\delta^2} = G(x)$$

equation (11) will go over into the differential equation:

$$f''(x) = -G(x) f(x).$$

Conversely, if we have a differential equation in the form

$$y'' = G(x)\, y \tag{13}$$

the Cauchy-Lipschitz existence theorem for solutions of differential equations† shows that, if $G(x)$ is continuous (not necessarily analytic) in the neighborhood of $a$, we can find a solution of (2) having the arbitrary boundary values

$$y(a) = c_1, \qquad y'(a) = c_2 \tag{14}$$

by regarding (13) as the limiting form of a difference equation. In the proof as usually presented (13) is replaced by the simultaneous pair of equations of the first order

$$p' = -G(x)\, y$$
$$y' = p$$

which are regarded as the limit of the difference equations

$$p_{n+1} - p_n = -G(a + n\delta)\, y_{n+1}\delta, \tag{15}$$
$$y_{n+1} - y_n = p_n\delta.$$

---

• *Cf.*, for example, Harnack, *Elemente der Differential- und Integralrechnung*, p. 56.

† See for instance Picard, *Traité d'analyse*, vol. 2, p. 291.

Here we must let

$$y_0 = c_1, \qquad p_0 = c_2.$$

By means of the two equations of (14), all $y$'s and $p$'s of higher order can be computed, and the Cauchy proof shows that

$$\lim_{n=\infty} y_n = y(x), \qquad \lim_{n=\infty} p_n = y'(x)$$

where $y(x)$ is a solution of (13).

The upper limit for $n\delta = |x - a|$ postulated in the classic proof is in general much too small. Picard and Painlevé* have shown that the Cauchy solution will be convergent throughout the whole interval $b_0 b_1$ ($b_0 < a < b_1$), in which $y(x)$ as well as $G(x)$ are continuous and that the convergence is uniform over every tract within this interval. Moreover, since we are dealing with a linear equation, the interval throughout which $y$ is continuous will be at least as great as the interval throughout which $G$ is continuous.†

It is well known‡ that no solution of (13) can have a multiple root at a point of the interval just considered. From this it follows that there cannot be two distinct solutions of (13) satisfying conditions (14), since their difference would have a multiple root at $a$. On the other hand it is clear that a solution of (2) which is not identically zero cannot have an infinite number of roots in a finite interval within and at the extremities of which $G$ is continuous. For, at the points of condensation of such roots, $y'$ as well as $y$ would be zero.

If we now suppose that $y(a)$, $y'(a)$, and $G(x)$ are continuous functions of a parameter $\lambda$ ($\lambda_0 < \lambda < \lambda_1$) then $y_n$ and $p_n$ are continuous functions of this parameter and, by the uniformity of the convergence, their limits $y(x)$ and $y'(x)$ are continuous functions of $(x, \lambda)$.

From the method in which $y_n$ and $p_n$ must be computed by means of (14) it is evident that the result will be the same if we eliminate $p$ between the two equations of (14), thus getting

$$y_{n+2} - 2y_{n+1} + y_n = -G(a + n\delta)\, y_{n+1}\, \delta^2. \tag{15}$$

If now we plot the ordinates $y_m$, $m = 0, 1, 2, \ldots n$, as in §1, using for the

* Picard, *Comptes Rendus*, vol. 128 (1899), p. 1363.
Painlevé, *Bull. de la Soc. Math. de France*, vol. 27 (1899), p. 150.

† *Cf.* Peano, *Math. Annalen*, vol. 32 (1888), p. 450.

‡ *Cf.* Sturm, *l. c.* p. 109.

corresponding abscissas $m\delta$, $m = 0, 1, \ldots n$, we shall get a broken line which, if $\delta$ be taken sufficiently small, will lie between the two curves

$$\left.\begin{aligned} y &= y(x) + \epsilon \\ y &= y(x) - \epsilon \end{aligned}\right\} \quad (\epsilon \text{ arbitrarily small})$$

and

where $x$ is supposed to be any point of an interval lying within the interval in which $y(x)$ is continuous.

Equation (15) belongs to the type of difference equations considered in §3, and it was then shown (*cf.* (8)) that if either $y_1/y_0$ or $G(\lambda)$ increases with $\lambda$, the other not decreasing, or both are functions of the parameter $\lambda$ which continually increase as $\lambda$ increases $y_n(\lambda)/y_{n-1}(\lambda)$ increases continually with $\lambda$ in any interval delimited by two successive roots of $y_n(\lambda) = 0$, and that the same is true of $\dfrac{y_n(\lambda) - y_{n-1}(\lambda)}{y_{n-1}(\lambda)\,\delta}$ so that we have the fundamental theorem:

*Under the restrictions imposed on $y'(a, \lambda)/y(a, \lambda)$ and $G(x, \lambda)$, namely, that either or both shall continually increase with $\lambda$, and neither decrease, the ratio $y'(x, \lambda)/y(x, \lambda)$ will be a continually increasing function of $\lambda$ in any interval delimited by two successive roots of $y(x, \lambda) = 0$.*

In the same way it can be shown that the roots of $y(x, \lambda) = 0$ *will continually increase with* $\lambda$.

The Oscillation Theorem for the differential equation can be arrived at in the same way by regarding it as a limiting form of the analogous theorem for the difference equation.

The equation (8) of §3 becomes, if we pass to the limit,

$$y(x, \lambda)\, y'_x(x, \lambda_1) - y(x, \lambda_1)\, y'_x(x, \lambda) =$$

$$\int_a^x [\, G(x,\lambda) - G(x,\lambda_1)\,]\, y(x,\lambda)\, y(x,\lambda_1)\, dx + y'_x(a,\lambda_1)\, y(a,\lambda) - y'_x(a,\lambda)\, y(a,\lambda_1).$$

This fundamental equation, which can be at once obtained from the differential equation itself, is the starting point of Sturm and Bôcher* in establishing the Comparison Theorems from which the other Sturmian Theorems are deduced.

YALE UNIVERSITY,
NEW HAVEN, CONNECTICUT.

* The Theorems of Oscillation of Sturm and Klein., *Bull. Am. Math. Soc.*, ser. 2, vol. 4 (1898), p. 295.

# SPACE OF CONSTANT CURVATURE.

By Frederick S. Woods.

As is well known, the investigation of the foundations of geometry has led, by three distinct methods, to the recognition of the fact that alongside of the Euclidean geometry there exist other systems, equally self consistent and equally capable of explaining the facts of experience. The first of these methods of research, that of Lobachevsky, Bolyai, and Gauss, proceeds upon the historic lines of Euclid. The second method, that of Klein, starts from Cayley's system of measurement with reference to a fundamental quadric and employs the principles of projective geometry. The third method, that of Riemann, consists essentially in the study of differential quadratic expressions and extends to space of any dimensions Gauss's idea of the curvature of a surface. The first two of these methods make less demand than does the third on special mathematical knowledge, and elementary expositions of them are readily accessible and widely known; but, as far as the present writer is aware, no elementary treatment of Riemann's ideas exists. Such a treatment seems desirable, however, not only because of the intrinsic importance of Riemann's methods, but also because his results have been sometimes misunderstood. The term "curvature of space" has led some to infer by a false analogy that Riemann's geometry is possible only on the assumption of a fourth dimension.*

The present article is an attempt to present Riemann's ideas in an elementary form. The basis of the article is Riemann's famous paper "Ueber die Hypothesen, welche der Geometrie zu Grunde liegen."† In addition, the writer has consulted, among others, the following papers:

| | |
|---|---|
| Dedekind-Weber: | Anmerkungen. Riemann's *Gesammelte Werke*, 1st ed., p. 384; 2nd ed., p. 405. |
| Helmholtz: | Ueber die thatsächlichen Grundlagen der Geometrie (1866). *Wissenschaftliche Abhandlungen*, vol. 2, no. 77, p. 610. |
| | Ueber die Thatsachen, die der Geometrie zum Grunde liegen (1868). *Ibid.*, no. 78, p. 618. |
| Beltrami: | Teoria fondamentale degli spazii di curvatura constante. *Annali di mat.*, ser. 2, vol. 2 (1868), p. 232. |

* This error has recently been repeated in an article by Professor E. S. Crawley on "Geometry: Ancient and Modern," published in the *Popular Science Monthly* for Jan., 1901, p. 257.

† *Gesammelte Werke*, 1st ed., p. 254; 2nd ed., p. 272.

Saggio di interpretazione della geometria non-Euclidea. *Giornale di mat.*, vol. 6 (1868), p. 284.

Newcomb: Elementary theorems relating to the geometry of a space of three dimensions and of uniform positive curvature . . . . . *Crelle*, vol. 83 (1877), p. 293.

Killing: Ueber zwei Raumformen mit constanter positiver Krümmung; *Crelle*, vol. 86 (1879), p. 72.

Die Rechnung in den Nicht-Euklidischen Raumformen. *Crelle*, vol. 89 (1880), p. 265.

Ueber die Clifford-Klein'schen Raumformen. *Math. Annalen*, vol. 39 (1891), p. 257.

*Einführung in die Grundlagen der Geometrie.* Padeborn, 1893.

Schur: Ueber die Deformation der Raüme constanten Riemann'schen Krümmungsmasses. *Math. Annalen*, vol. 27 (1886), p. 163.

Ueber den Zusammenhang der Raüme constanten Riemann'schen Krümmungsmasses mit den projectiven Raümen. *Ibid.*, vol. 27 (1886), p. 537.

Klein: *Autographirte Vorlesungen über Nicht-Euklidische Geometrie.* Göttingen, 1892.

Zur Nicht-Euklidischen Geometrie. *Math. Annalen*, vol. 37 (1890), p. 544.

*Lectures on Mathematics*, Lecture 11, p. 85, New York, 1894.

Lie-Engel: *Theorie der Transformationsgruppen*, vol. 3, p. 393 ff. Leipzig, 1893.

Bianchi: *Differentialgeometrie*, German translation by Max Lukat, p. 563 ff.

**1. The First Two Hypotheses.** Riemann's investigations are concerned with the properties of extents (*Mannigfaltigkeiten*) of any number of dimensions, in which an element may be determined by means of coordinates. Such extents are now commonly called "spaces of $n$ dimensions"; but we shall restrict the word "space", in this article, to an extent of three dimensions, while "surface" shall mean an extent of two dimensions, and "line" an extent of one dimension. We have then two problems. First, we have to develop a system of geometry from assumed hypotheses and, secondly, we have to bring our results to the test of experience and inquire how far they are in accord with natural phenomena.

These two problems need to be sharply distinguished. The first is theoretical and exact; the second is empirical and subject to all the uncertainty which belongs to physical observations. In attacking the first problem, the investigator is free to choose his hypotheses as he will, provided only they are not contradictory, and to draw what conclusions he may from them. We shall accordingly make the two following hypotheses which we will first state and afterwards discuss.

First Hypothesis. *The space to be considered shall be a continuum of three dimensions, in which a point may be determined by three independent real quantities, $x_1$, $x_2$, $x_3$.*

SECOND HYPOTHESIS. *The length of a line shall be determined by means of a line element given by the equation*

$$ds = \sqrt{\Sigma a_{ik}\, dx_i\, dx_k}\,, \qquad (a_{ki} = a_{ik};\ i, k = 1, 2, 3)$$

*where the $a_{ik}$ are functions of $x_1$, $x_2$, $x_3$, which together with their differential coefficients of the first four orders are finite and continuous, and where the expression under the radical sign is positive for all real values of $x_1$, $x_2$, $x_3$ and $dx_1$, $dx_2$, $dx_3$, provided that $x_1$, $x_2$, $x_3$ correspond to points of the space and that $dx_1$, $dx_2$, $dx_3$ are not all zero.*

*Discussion of the First Hypothesis.* In a purely analytic sense, a "point" is a set of three values $(x_1, x_2, x_3)$, and "space" is a collection, or assemblage, of such points, which together form a continuum of three dimensions. By the requirement that space be a continuum, the following two properties are asserted of space. First, if $(a_1, a_2, a_3)$ is any point of space, then $(a_1 + h_1, a_2 + h_2, a_3 + h_3)$ is also a point of space, whatever may be the values of $h_1$, $h_2$, and $h_3$, provided that no one of them exceeds numerically a certain value $h$, depending in general on the point $(a_1, a_2, a_3)$. Secondly, if $(a_1, a_2, a_3)$ and $(b_1, b_2, b_3)$ are any two points of space, it is possible to connect them by a continuous curve lying in the space.

Here it is immaterial whether $x_1$, $x_2$, and $x_3$ be allowed to take all possible values, or whether they be subject to certain restrictions, provided only the above two conditions be met.

We assume for the present that to every point of space, or at least to every point of the portion of space which we are studying, corresponds one set of coordinates $(x_1, x_2, x_3)$, and conversely. We can later proceed to all portions of space, in case this condition is not everywhere fulfilled, by the usual methods of analysis.

Our space is assumed to be three dimensional. By this it is not meant to deny the possibility of the space's lying in an extent of higher dimensions, but simply to assert that the discussion has nothing whatever to do with these dimensions. The validity of the results obtained is neither strengthened nor weakened by the assumption of a fourth dimension.

*Discussion of the Second Hypothesis.* Within the space thus defined we may pick out at pleasure one dimensional extents or lines. We shall restrict ourselves here to lines which may be expressed analytically by the equations

$$x_1 = f_1(t), \qquad x_2 = f_2(t), \qquad x_3 = f_3(t),$$

where $t$ is an arbitrary parameter, and where, for the values of $t$ which come into consideration, $f_1$, $f_2$, and $f_3$ are continuous functions possessing continuous first derivatives which do not vanish simultaneously. Consider now a portion of such a line corresponding to values of $t$ lying between, and including, the values $t_0$ and $T$, and let this portion be broken up into $n$ segments, to the extremities of which correspond the values $t_0, t_1, t_2, \ldots, t_{n-1}, T$. If now $(x_1, x_2, x_3)$ are the coordinates of the extremity $(t = t_i)$ of any segment, the coordinates of the other extremity $(t = t_{i+1})$ will be, except for infinitesimals of higher order, $(x_1 + dx_1, x_2 + dx_2, x_3 + dx_3)$, where

$$dx_1 = f_1'(t)dt, \quad dx_2 = f_2'(t)dt, \quad dx_3 = f_3'(t)dt, \text{ and } dt = t_{i+1} - t_i.$$

We introduce now arbitrarily as the definition of an *element of length* a function

$$\phi(x_1, x_2, x_3; dx_1, dx_2, dx_3)$$

which has the following two properties. First, it shall become infinitesimal with $dx_1$, $dx_2$, $dx_3$ and consequently with $dt$, and, secondly, the sum of the $n$ values of this function computed for the $n$ segments of the line considered shall approach a limit as $n$ is indefinitely increased and at the same time the $n$ quantities $t_{i+1} - t_i$ are made to approach zero: the limit to be independent of the manner in which the extremities of the segments are chosen. *This limit is defined as the length of the portion of the line in question.*

These conditions are met in particular by the assumption of the line element as in the second hypothesis, in accordance with which we have

$$\phi = \sqrt{\Sigma a_{ik}\, dx_i\, dx_k}.$$

The length of the line is then given by the integral

$$l = \int_{t_0}^{T} \sqrt{\Sigma a_{ik} f_i'(t) f_k'(t)}\; dt.$$

This form of the line element has for our present purpose the decided advantage of including as special cases all the forms occurring in Euclidean geometry, which correspond to different systems of coordinates (Cartesian, polar, cylindrical, etc.); but it is by no means the only conceivable one. Thus we might, as pointed out by Riemann, employ the positive fourth root of a homogeneous biquadratic expression in $dx_1$, $dx_2$, $dx_3$. Further, the restriction that the expression $\Sigma a_{ik}\, dx_i\, dx_k$ should be always positive for allowable real values

of the quantities involved may be dropped,* but in that case we should have certain real lines with zero or even with imaginary length. Such a geometry would not explain the facts of experience and is therefore not considered here.

**2. Geodesic Lines and Surfaces.** We proceed now to find the equations of the shortest line between two points; or, in other words, to determine the condition that the integral

$$I = \int_{t_0}^{t_1} \sqrt{\Sigma a_{ik} \frac{dx_i}{dt} \frac{dx_k}{dt}}\, dt$$

shall be a minimum. The Calculus of Variations gives, as a necessary condition that the integral

$$I = \int_{t_0}^{t_1} F(x, y, z, x', y', z')\, dt$$

be a minimum, the Eulerian equations

$$\frac{d}{dt} F_{x'} - F_x = 0, \qquad \frac{d}{dt} F_{y'} - F_y = 0, \qquad \frac{d}{dt} F_{z'} - F_z = 0.$$

In the present case, these become

$$\frac{d}{dt} \frac{1}{\sqrt{R_t}} \left( a_{l1} \frac{dx_1}{dt} + a_{l2} \frac{dx_2}{dt} + a_{l3} \frac{dx_3}{dt} \right) = \tfrac{1}{2} \frac{1}{\sqrt{R_t}} \sum_{ik} \frac{\partial a_{ik}}{\partial x_l} \frac{dx_i}{dt} \frac{dx_k}{dt},$$

where $l = 1, 2, 3$ and $R_t = \Sigma a_{ik} \frac{dx_i}{dt} \frac{dx_k}{dt}$. The integral $I$ is invariant of the parameter $t$. We introduce the parameter $s$ defined by the equation

$$s = \int_{t_0}^{t} \sqrt{\Sigma a_{ik} \frac{dx_i}{dt} \frac{dx_k}{dt}}\, dt,$$

*i. e.* the length $s$ of the portion of the geodesic line between an assumed fixed point and a variable point. Since

$$\int_{t_0}^{t} \sqrt{R_t}\, dt = \int_{s_0}^{s} \sqrt{R_s}\, ds = s,$$

* Compare for example Bianchi's paper: "Sulle forme differenziali quadratiche indefinite," *Atti della R. Accademia dei Lincei*, Memorie, ser. 4, vol. 5 (1888), p. 539, as well as the writer's article: Ueber Pseudominimalflächen, Göttingen 1895, wherein are developed properties of certain surfaces in a space for which the line element is $ds = \sqrt{dx^2 + dy^2 - dz^2}$.

it follows that

$$\Sigma a_{ik} \frac{dx_i}{ds} \frac{dx_k}{ds} = 1.$$

Hence the necessary conditions for a minimum, in terms of the parameter $s$, are

$$\left.\begin{aligned}
\frac{d}{ds}\left[a_{11}\frac{dx_1}{ds} + a_{12}\frac{dx_2}{ds} + a_{13}\frac{dx_3}{ds}\right] &= \tfrac{1}{2}\sum \frac{\partial a_{ik}}{\partial x_1}\frac{dx_i}{ds}\frac{dx_k}{ds}\,, \\
\frac{d}{ds}\left[a_{12}\frac{dx_1}{ds} + a_{22}\frac{dx_2}{ds} + a_{23}\frac{dx_3}{ds}\right] &= \tfrac{1}{2}\sum \frac{\partial a_{ik}}{\partial x_2}\frac{dx_i}{ds}\frac{dx_k}{ds}\,, \\
\frac{d}{ds}\left[a_{13}\frac{dx_1}{ds} + a_{23}\frac{dx_2}{ds} + a_{33}\frac{dx_3}{ds}\right] &= \tfrac{1}{2}\sum \frac{\partial a_{ik}}{\partial x_3}\frac{dx_i}{ds}\frac{dx_k}{ds}\,.
\end{aligned}\right\} \qquad (A)$$

Conversely, these conditions are sufficient if $l$ is not too great. More precisely: Let $(a_1, a_2, a_3)$ be any point of space, and $(x_1, x_2, x_3)$ any second point such that $|x_i - a_i|$ does not exceed a suitably chosen positive quantity, $h$. Then the equations $(A)$ admit one and only one solution which passes through the points $(a)$ and $(x)$ and has all of its points lying in the region $|x_i - a_i| < h$, and for the corresponding curve the integral $l$ has a smaller value than for any other curve joining the points $(a)$ and $(x)$.

We take the equations $(A)$ accordingly, as the defining equations of the *geodesic lines*. In accordance with the theory of differential equations, it is always possible to find one and only one solution of these equations which takes on at an arbitrary point $(x_1, x_2, x_3)$ any arbitrarily assigned values (not all zero) of the differential coefficients $dx_1/ds$, $dx_2/ds$, $dx_3/ds$. If these differential coefficients satisfy initially the relation

$$\Sigma a_{ik} \frac{dx_i}{ds} \frac{dx_k}{ds} = 1,$$

this relation will be fulfilled for all values of $s$.

The geodesic lines which radiate from a point are hence distinguished from each other by the ratios of the values of the differential coefficients and these may consequently be regarded as fixing the direction of each line; the direction being, broadly, that property of a geodesic line which distinguishes it from all others through the same point.

It will be convenient to denote $dx_i/ds$ by $\xi_i$, and to define the triple consisting of the three direction quantities $\xi_1, \xi_2, \xi_3$ as the *direction* $(\xi_1, \xi_2, \xi_3)$. These quantities satisfy the relation

$$\Sigma a_{ik} \xi_i \xi_k = 1. \qquad (B)$$

Our results thus far are these:

THEOREM 1. *A geodesic line is completely determined by means of any point on it and any second point not too far distant from the first.*

THEOREM 2. *A geodesic line is completely determined by means of any point on it and its direction at that point.*

THEOREM 3. *Through any point of space pass $\infty^2$ geodesic lines, each of which is completely determined by means of its initial direction.*

Take now two geodesic lines extending from the same point with the directions $(a_1, a_2, a_3)$ and $(\beta_1, \beta_2, \beta_3)$ and consider the system of geodesic lines which radiate from their point of intersection with the directions $(\xi_1, \xi_2, \xi_3)$, where

$$\begin{aligned} \xi_1 &= \lambda a_1 + \mu \beta_1, \\ \xi_2 &= \lambda a_2 + \mu \beta_2, \\ \xi_3 &= \lambda a_3 + \mu \beta_3, \end{aligned}$$

and, by virtue of the relation $(B)$, $\lambda$ and $\mu$ are connected by the relation

$$\lambda^2 + \mu^2 + 2\lambda\mu \Sigma a_{ik} a_i \beta_k = 1.$$

Such a system we call a *pencil of geodesic lines*, and the common point the *vertex*. This pencil corresponds to the totality of straight lines in elementary Euclidean geometry, all of which go through a common point and lie in the same plane. Any two lines of a pencil define the pencil. Any line of the pencil is fixed by means of a single parameter, which may be conveniently chosen by introducing the conception of an *angle*. *The angle $\theta$ between two intersecting geodesic lines, with the directions $(\xi_1, \xi_2, \xi_3)$ and $(\xi_1', \xi_2', \xi_3')$, shall be defined by the relation*

$$\cos\theta = \Sigma a_{ik} \xi_i \xi_k',$$

*where the $a_{ik}$ are computed for the point of intersection.*

The angle $\theta$ is always real, since we are considering only real geodesic lines. This may be most conveniently seen by dropping for the moment the condition that $(\xi_1, \xi_2, \xi_3)$ should satisfy relation $(B)$. We may then interpret these quantities as homogeneous point coordinates in a Euclidean plane. If now $(\xi_1', \xi_2', \xi_3')$ and $(\xi_1'', \xi_2'', \xi_3'')$ are two points of this plane, any point of the line connecting them is given by the formula:

$$\xi_1 = \lambda \xi_1' + \mu \xi_2'',$$
$$\xi_2 = \lambda \xi_1' + \mu \xi_2'',$$
$$\xi_3 = \lambda \xi_1' + \mu \xi_2'';$$

and the points in which this line intersects the conic section

$$\Sigma a_{ik} \xi_i \xi_k = 0$$

correspond to values of $\lambda : \mu$ given by the equation:

$$\lambda^2 \Sigma a_{ik} \xi_i' \xi_k' + 2\lambda\mu \Sigma a_{ik} \xi_i' \xi_k'' + \mu^2 \Sigma a_{ik} \xi_i'' \xi_k'' = 0.$$

But the above conic has no real points of intersection by virtue of our second hypothesis. Therefore the roots of the quadratic in $\lambda : \mu$ are imaginary. Therefore its coefficients satisfy the relation:

$$[\Sigma a_{ik} \xi_i' \xi_k'][\Sigma a_{ik} \xi_i'' \xi_k''] > [\Sigma a_{ik} \xi_i' \xi_k'']^2.$$

Introducing now the condition $(B)$ we have

$$|\Sigma a_{ik} \xi_i' \xi_k''| < 1.$$

The limiting cases $\Sigma a_{ik} \xi_i' \xi_k'' = \pm 1$ can occur only when

$$\xi_1' : \xi_2' : \xi_3' = \xi_1'' : \xi_2'' : \xi_3''.$$

Since the signs of $\xi_1$, $\xi_2$, and $\xi_3$ may be changed without altering the geodesic line corresponding to them (for such a change of sign may be made by altering the sign of $ds$, that is, of the direction in which the length is measured), it follows that $\cos\theta$ is ambiguous in sign. This ambiguity is the same as that met with in the angle between straight lines in Euclidean geometry and may be removed by familiar conventions.

If $\cos\theta = 0$: *i. e.* if

$$\Sigma a_{ik} \xi_i \xi_k' = 0$$

the two lines are said to be *perpendicular*.

We may now readily prove the following theorems:

THEOREM 4. *In any pencil of geodesic lines there exists one and only one line perpendicular to any given line of the pencil.*

THEOREM 5. *If a geodesic line through the vertex of a pencil is perpendicular to two lines of the pencil, it is perpendicular to all lines of the pencil.*

THEOREM 6. *Through the vertex of a pencil of geodesic lines passes one and only one geodesic line perpendicular to all lines of the pencil.*

THEOREM 7. *The geodesic lines perpendicular to a given geodesic line at a given point form a pencil.*

*Proof of Theorem* 4. Let a pencil be given with its vertex at $O$ and let $OA$ and $OB$ be any two lines of the pencil with directions $(a_1, a_2, a_3)$ and $(\beta_1, \beta_2, \beta_3)$ respectively, making the angle $\omega$ with each other. Then

$$\cos\omega = \Sigma a_{ik}\, a_i \beta_k.$$

Let $OM$ be any line of the pencil defined by $OA$ and $OB$; then the direction of $OM$ is given by the equations

$$\xi_i = \lambda a_i + \mu \beta_i,$$

where $$\lambda^2 + \mu^2 + 2\lambda\mu\cos\omega = 1.$$

The angle $\theta$ between $OM$ and $OA$ is given by the formula

$$\cos\theta = \Sigma a_{ik}\, a_i\, (\lambda a_k + \mu\beta_k) = \lambda + \mu\cos\omega.$$

The condition that $OM$ be perpendicular to $OA$ is that $\cos\theta = 0$; hence we have as the necessary and sufficient condition for perpendicularity:

$$\lambda = \pm\cot\omega, \quad \mu = \mp\csc\omega.$$

The direction of $OM$ is thus restricted to a single direction and its opposite, and hence there is one and only one line in the pencil perpendicular to $OA$.

*Proof of Theorem* 5. Let $OC$ be a geodesic line through $OA$ perpendicular to the line $OA$ and $OB$ of the pencil. The direction $(\gamma_1, \gamma_2, \gamma_3)$ of $OC$ satisfies, then, the conditions

$$\Sigma a_{ik}\, a_i\, \gamma_k = 0, \qquad \Sigma a_{ik}\, \beta_i\, \gamma_k = 0,$$

and consequently the condition

$$\Sigma a_{ik}\, (\lambda a_i + \mu\beta_i)\, \gamma_k = 0$$

for all values of $\lambda$ and $\mu$. Hence $OC$ is perpendicular to every line of the pencil.

*Proof of Theorem* 6. By the equations

$$\Sigma a_{ik}\, a_i\, \gamma_k = 0, \qquad \Sigma a_{ik}\, \beta_i\, \gamma_k = 0,$$

the values of the ratios of $\gamma_1 : \gamma_2 : \gamma_3$ are uniquely determined, and by aid of $(B)$ the values of these quantities are determined except for sign. Hence it

is always possible to determine one and only one perpendicular to two intersecting geodesic lines at their point of intersection. By Theorem 5, this line is perpendicular to every line of the pencil.

*Proof of Theorem* 7. If $OC$ with direction $(\gamma_1, \gamma_2, \gamma_3)$ be given, the condition that a geodesic line through $O$ should be perpendicular to $OC$ is

$$\Sigma a_{ik} \gamma_i \xi_k = 0.$$

If now $(\alpha_1, \alpha_2, \alpha_3)$ and $(\beta_1, \beta_2, \beta_3)$ are two sets of values which satisfy this equation, it will be satisfied by all values of the form

$$\xi_i = \lambda \alpha_i + \mu \beta_i$$

and by no other values. This follows from the fact that the equation is linear in the unknown quantities $\xi_i$. Hence the geodesic lines through $O$ perpendicular to $OC$ form a pencil.

*Geodesic Surface.* If we define a pencil by means of two perpendicular geodesic lines $OA$ and $OB$, and denote as before by $\theta$ the angle (properly chosen) between $OA$ and any line of the pencil $OM$, we have

$$\lambda = \cos\theta, \qquad \mu = \sin\theta;$$

so that the direction of $OM$ is given by

$$\xi_i = \alpha_i \cos\theta + \beta_i \sin\theta,$$

where $\theta$ is arbitrary and $\Sigma a_{ik} \alpha_i \beta_k = 0$.

If $P(x_1, x_2, x_3)$ be any point on the line $OM$ and $r$ be the length of this line from $O$ to $P$, the coordinates of $P$ are determined by integrating the differential equations of the geodesic line, choosing the solution which at the point $x_i$ has the direction $\xi_i$, and substituting $r$ for $s$. We have then

$$x_i = f_i(\xi_1, \xi_2, \xi_3, r) = \phi_i(\theta, r),$$

the functions $\phi_i(\theta, r)$ being continuous, together with their partial derivatives of the first and second orders. By giving to $\theta$ and $r$ all possible values, all points of all lines belonging to the pencil are determined. These points, dependent upon two parameters, form an extent of two dimensions which we call a *geodesic surface*. *A geodesic surface is then defined as a pencil of geodesic lines.* By the definition, the vertex of the pencil forms a unique point of the surface. Whether or not any point of the surface may play this important rôle

is at present undecided. We cannot in fact assume that the geodesic surface contains any geodesic lines of space other than those of the pencil.

**3. Curvature.** *a. The Curvature of a Surface.* Let there be given any surface

$$x_i = \phi_i(u, v), \qquad (i = 1, 2, 3)$$

where $u$ and $v$ are independent variables and the $\phi_i$ are singled valued continuous functions with continuous differential coefficients of the first two orders. The line element of any line on the surface is found from the line element of space. There results

$$ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2,$$

where

$$E = \Sigma a_{ik} \frac{\partial x_i}{\partial u} \frac{\partial x_k}{\partial u},$$

$$F = \Sigma a_{ik} \frac{\partial x_i}{\partial u} \frac{\partial x_k}{\partial v},$$

$$G = \Sigma a_{ik} \frac{\partial x_i}{\partial v} \frac{\partial x_k}{\partial v}.$$

The properties of the surface belong to two distinct classes. The properties of the first class are those which may be deduced from a study of the line element just written, where $E$, $F$, $G$ are treated as functions of $u$ and $v$, and no reference is had to their dependence upon $x_1, x_2, x_3$; while those of the second class depend upon the relations between $u$, $v$ and $x_1, x_2, x_3$. For example, if the line element of the surface is

$$ds^2 = du^2 + dv^2,$$

the surface may be a plane, a cylinder, or any developable surface, if it lies in Euclidean space, and may be something quite different in a non-Euclidean space, and the properties of the first class are those common to all these surfaces. The distinction between the two classes of properties is often lost sight of in elementary presentations of the theory of surfaces, but it is of fundadamental importance to us, since the properties of the first class maintain their significance whatever may be the nature of the space in which the surface lies. We may accordingly use the results of the ordinary theory of surfaces, in so far as they concern these properties. To obtain the properties of the first class, we assume the line element of the surface, and define length, geodesic

lines on the surface, and angle in a manner similar to that just employed for space. Thus the equations of the geodesic lines are

$$\frac{d}{ds}\left[E\frac{du}{ds}+F\frac{dv}{ds}\right]=\tfrac{1}{2}\left[\frac{\partial E}{\partial u}\left(\frac{du}{ds}\right)^2+2\frac{\partial F}{\partial u}\frac{du}{ds}\frac{dv}{ds}+\frac{\partial G}{\partial u}\left(\frac{dv}{ds}\right)^2\right]$$

$$\frac{d}{ds}\left[F\frac{du}{ds}+G\frac{dv}{ds}\right]=\tfrac{1}{2}\left[\frac{\partial E}{\partial v}\left(\frac{du}{ds}\right)^2+2\frac{\partial F}{\partial v}\frac{du}{ds}\frac{dv}{ds}+\frac{\partial G}{\partial v}\left(\frac{dv}{ds}\right)^2\right]$$

and the angle between two geodesics is defined by the formula

$$\cos\theta =$$

$$E\left(\frac{du}{ds}\right)'\left(\frac{du}{ds}\right)''+F\left[\left(\frac{du}{ds}\right)'\left(\frac{dv}{ds}\right)''+\left(\frac{dv}{ds}\right)'\left(\frac{du}{ds}\right)''\right]+G\left(\frac{dv}{ds}\right)'\left(\frac{dv}{ds}\right)''.$$

It should be noticed that, if the surface is considered as lying in a given space, the geodesic lines on the surface are not in general geodesic lines of the space, but lengths and angles are the same whether measured on the surface or in space.

The properties which may be deduced for a surface by the discussion of its line element depend largely upon the quantity called by Gauss the measure of curvature, or simply the *curvature*, of the surface. Under that term we understand a quantity $K$, defined by the relation

$$K=\frac{1}{2\sqrt{EG-F^2}}\left\{\frac{\partial}{\partial u}\left[\frac{F}{E\sqrt{EG-F^2}}\frac{\partial E}{\partial v}-\frac{1}{\sqrt{EG-F^2}}\frac{\partial G}{\partial u}\right]\right.$$
$$\left.+\frac{\partial}{\partial v}\left[\frac{2}{\sqrt{EG-F^2}}\frac{\partial F}{\partial u}-\frac{1}{\sqrt{EG-F^2}}\frac{\partial E}{\partial v}-\frac{F}{E\sqrt{EG-F^2}}\frac{\partial E}{\partial u}\right]\right\}.$$

With the geometric interpretation of the curvature as usually given on the hypothesis that the surface lies in Euclidean space we have nothing to do. For us, the "curvature" means simply the above expression, which is fully determined when the line element of the surface is given, and may be shown to be an invariant of the surface, that is, independent of the coordinates chosen to define the surface. When $K$ is the same for all points of the surface, the surface is said to be one of *constant curvature*. The importance of the curvature lies in the two theorems:

*A necessary condition that two portions of surfaces may be brought into point for point correspondence with preservation of distance (and hence of angle) is that they have the same curvature at corresponding points.*

*If the two portions of surfaces are of constant curvature, the condition is also sufficient.*

It follows that on any surface of constant curvature there exists a geometry of geodesic lines which is independent of the portion of the surface considered.

Beltrami* has shown that this geometry assumes different distinctive features according as the constant curvature is positive, negative, or zero.

*b. The Curvature of Space.* As we have seen, the coordinates of a point on a geodesic surface may be taken as $(r, \theta)$. The element of arc of a geodesic surface is then

$$ds^2 = E\,dr^2 + 2F\,dr\,d\theta + G\,d\theta^2,$$

where it is possible to determine more exactly the values of the coefficients. In the first place, if $\theta = \text{const.}$, $ds = dr$ for all values of $r$, no matter what value may have been assigned to $\theta$; hence $E = 1$. Secondly, the lines $r = s$, $\theta = \text{const.}$ are geodesic lines in space and hence geodesic lines on the surface. Substituting the values $E = 1$, $\frac{d\theta}{ds} = 0$ in the equations of the geodesic lines we have

$$\frac{\partial F}{\partial r} = 0,$$

and therefore $F$ is a function of $\theta$ only. When $r = 0$ the quantities $x_1, x_2$, and $x_3$ reduce to the coordinates of the vertex of the generating pencil, no matter what the value of $\theta$. Hence

$$\left(\frac{\partial x_1}{\partial \theta}\right)_{r=0} = 0, \qquad \left(\frac{\partial x_2}{\partial \theta}\right)_{r=0} = 0, \qquad \left(\frac{\partial x_3}{\partial \theta}\right)_{r=0} = 0.$$

From the formula

$$F = \Sigma a_{ik} \frac{\partial x_i}{\partial r} \frac{\partial x_k}{\partial \theta},$$

it follows that $(F)_{r=0} = 0$, and hence, since $F$ is a function of $\theta$ only, $F \equiv 0$. The line element of the geodesic surface is therefore

$$ds^2 = dr^2 + G\,d\theta^2.$$

---

* Saggio di interpretationi della geometria non-Euclidea, *Giornale di matematiche*, vol. 6, (1868), p. 284, *cf.* also, Bianchi, *Differentialgeometrie*, p. 418 ff.

The expression of the curvature reduces in this case to

$$K = -\frac{1}{\sqrt{G}}\frac{\partial^2\sqrt{G}}{\partial r^2}.$$

The curvature is consequently a function not only of $r$ and $\theta$, the coordinates of a point on the surface, but also in general of the coordinates of the vertex $O$ of the generating pencil and of the directions $(a_1, a_2, a_3)$ and $(\beta_1, \beta_2, \beta_3)$ which fix the pencil, since the functions $\phi_i(r, \theta)$ depend upon these latter quantities. In particular, if $K_0$ is the curvature at $O$, it may be shown that

$$K_0 = -\tfrac{1}{2}\frac{\sum_{iklm}(ik, lm)_0\,(a_i\beta_k - a_k\beta_i)(a_l\beta_m - a_m\beta_l)}{\sum_{iklm}(a_{il}a_{km} - a_{im}a_{kl})_0\,(a_i\beta_k - a_k\beta_i)(a_l\beta_m - a_m\beta_l)},$$

where

$$(ik, lm) =$$

$$\frac{\partial^2 a_{il}}{\partial x_k \partial x_m} + \frac{\partial^2 a_{km}}{\partial x_i \partial x_l} - \frac{\partial^2 a_{im}}{\partial x_k \partial x_l} - \frac{\partial^2 a_{kl}}{\partial x_i \partial x_m} + \tfrac{1}{2}\sum_{st}\left(p_{skm}p_{til} - p_{sim}p_{tkl}\right)\frac{A_{st}}{a},$$

$$p_{\lambda\mu\nu} = \frac{\partial a_{\lambda\mu}}{\partial x_\nu} + \frac{\partial a_{\lambda\nu}}{\partial x_\mu} - \frac{\partial a_{\mu\nu}}{\partial x_\lambda},$$

$A_{st}$ is the minor corresponding to $a_{st}$ in the determinant

$$a = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix},$$

the subscript 0 denotes that the expressions to which it is affixed are computed for the point $O$, and the summation is to be extended over all combinations of the value-pairs $i, k$ and $l, m$, for which $i < k$, $l < m$.*

The quantity $K_0$ is called by Riemann the *measure of curvature of the space* in a given point and in a given surface direction. This curvature varies in general from point to point in space, and also in a given point is different for the different geodesic surfaces which pass through the point. In case $K_0$ is constant, the space is said to be one of *constant curvature*. *A space of constant curvature is accordingly one in which the curvatures of all geodesic surfaces computed for the vertices of the generating pencils are the same.*

* *Cf.* Dedekind-Weber, *l. c.* also Schur, *Math. Annalen*, vol. 27 (1886), pp. 542–551. A shorter derivation involving the use of differential invariants and Christoffel's symbolic expressions is given by Bianchi, *l. c.* p. 571 ff.

From the expression for $K_0$ it follows that the necessary and sufficient condition for a space of constant curvature is that the quantities $a_{ik}$ should satisfy identically the relations

$$(ik, lm) = K_0(a_{il}a_{km} - a_{im}a_{kl}),$$

where $K_0$ is a constant.

**4. The Third Hypothesis.** We shall define a *displacement* as *any transformation by which a continuous portion of space is brought into a continuous point for point correspondence either with itself or with some other portion of space in such a manner that the lengths of corresponding portions of lines are the same.*

Let $S$ be any portion of space in which the coordinates of any point $P$ are $(x_1, x_2, x_3)$, and let $S'$ be a portion of space in which the coordinates of a point $P'$ are $(x'_1, x'_2, x'_3)$. Let the line element in $S$ be denoted by

$$ds = \sqrt{\Sigma a_{ik}\, dx_i\, dx_k}\,,$$

and the line element in $S'$ by

$$ds' = \sqrt{\Sigma a'_{ik}\, dx'_i\, dx'_k}\,,$$

where $a'_{ik}$ denotes the value of $a_{ik}$ for $(x'_1, x'_2, x'_3)$. In order that $S$ may be displaced into $S'$, it is necessary and sufficient that

$$ds = ds'$$

by virtue of relations of the form

$$x_i = \phi_i(x'_1, x'_2, x'_3), \qquad (i = 1, 2, 3)$$

where the $\phi_i$ are continuous functions of $(x'_1, x'_2, x'_3)$, possessing continuous first derivatives, and establishing a one-to-one relation between the points of $S$ and $S'$.

From this follows

$$a'_{ik} = \sum_{lm} a_{lm} \frac{\partial \phi_l}{\partial x'_i} \frac{\partial \phi_m}{\partial x'_k}.$$

*Any displacement transforms geodesic lines into geodesic lines, and geodesic surfaces into geodesic surfaces.*

The first part of this theorem follows at once from the definition of geodesic lines. To prove the second part, we notice that

$$\frac{dx_i}{ds} = \sum_l \frac{\partial \phi_i}{\partial x_l'} \frac{dx_l'}{ds'} \frac{ds'}{ds} = \sum_l \frac{\partial \phi_i}{\partial x_l'} \frac{dx_l'}{ds'}.$$

Hence, if $(a_1, a_2, a_3)$ is the direction of a geodesic line at a point $O$, and $(a_1', a_2', a_3')$ the direction of the corresponding line at $O'$,

$$a_i = \sum_l \left(\frac{\partial \phi_i}{\partial x_l'}\right)_0 a_l'.$$

Since this relation is linear, a pencil of geodesic lines with its vertex at $O$ is transformed into a pencil of geodesics with the vertex at $O'$.

*Under any displacement the angle between two geodesic lines is equal to the corresponding angle between the two corresponding geodesic lines.*

For let $OA$ and $OB$ be two geodesic lines intersecting at $O$ with the directions $(a_1, a_2, a_3)$ and $(\beta_1, \beta_2, \beta_3)$ respectively, and let $O'B'$ and $O'A'$ be the corresponding lines with directions $(a_1', a_2', a_3')$ and $(\beta_1', \beta_2', \beta_3')$. Then if $\theta$ is the angle between $OA$ and $OB$ and $\theta'$ the angle between $O'A'$ and $O'B'$,

$$\begin{aligned} \cos\theta &= \sum_{ik} a_{ik} a_i \beta_k \\ &= \sum_{lm} \left\{ \sum_{ik} a_{ik} \frac{\partial \phi_i}{\partial x_l'} \frac{\partial \phi_k}{\partial x_m'} \right\} a_l' a_m' \\ &= \sum_{lm} a_{lm}' a_l' a_m' = \cos\theta'. \end{aligned}$$

To remove the ambiguity involved in the determination of an angle from its cosine, and at the same time to define corresponding angles, we may consider first the case in which $OB$ coincides with $OA$. Then $O'B'$ coincides with $O'A'$ and we may take $\theta = \theta' = 0$. Let now $OB$ change its position continuously, the point $O$ being fixed: $\theta$ then varies continuously, attaining a final value which we call the angle $AOB$. At the same time $O'B'$ changes its position continuously, always corresponding to $OB$, and $\theta'$ varies continuously from zero to a final value $A'O'B'$, which is the angle corresponding to $AOB$. Since at all times $\cos\theta = \cos\theta'$, and at the outset $\theta = \theta'$, it follows that $AOB = A'O'B'$.

The possibility of the reversal of the direction of the angle is not excluded, and it seems unnecessary at this point to introduce a distinction between displacements which do not reverse the direction of the angle and those which do.

The existence of displacements is by no means necessary in a general space. This is to be seen by comparing the similar problem in connection with surfaces. A displacement on a surface may be carried out by bending one portion of the surface upon another portion without tearing or stretch-

ing. It is clear that surfaces exhibit widely different possibilities in regard to such bending. It may also be seen that the displacement of a portion of a surface and of the entire surface, are different problems. Take, for example, the right circular cylinder in Euclidean space. If a point on the cylinder is fixed, the cylinder as a whole is immovable, but a restricted portion of the cylinder around the fixed point is easily bent upon itself.

The existence of displacements in space is consequently the subject of a new hypothesis.

THIRD HYPOTHESIS. *If $P$ is any point of space, it shall be possible to displace a restricted portion of space surrounding $P$ upon itself in such a manner that any two geodesic lines through $P$ shall correspond to any other two geodesic lines through $P$, provided only that the two latter lines make the same angle with each other as do the former lines.*

By such a displacement, the point $P$ remains fixed and any two geodesic pencils with their vertices at $P$ may be brought into point for point correspondence with preservation of distance and angle. Hence (*cf.* §3, *a*), the geodesic surfaces formed by the two pencils must have the same curvature at corresponding points and in particular at $P$. The quantity $K_0$ of the previous article must therefore be independent of $(a_1, a_2, a_3)$ and $(\beta_1, \beta_2, \beta_3)$. Now Schur has proved (*l. c.*) that when the curvature of space is constant at each point it does not change as we pass from point to point. Hence $K_0$ must also be independent of the coordinates of $O$. We have therefore the theorem:

*In order that the third hypothesis shall be met, it is necessary that the space be one of constant curvature.*

In the next paragraph we shall show that in a space of constant curvature the hypothesis is always met and that furthermore there exist displacements by which any point $P$ may be transformed into any other point $P'$.

**5. Normal Form of the Line Element.** Let the line element

$$ds^2 = \Sigma a_{ik}\, dx_i\, dx_k \tag{1}$$

belong to a space of constant curvature $k^2$, where $k$ may be zero, real, or pure imaginary. It is then always possible by a change of coordinates to reduce the line element to Riemann's normal form

$$ds^2 = \frac{dy_1^2 + dy_2^2 + dy_3^2}{\left[1 + \frac{k^2}{4}(y_1^2 + y_2^2 + y_3^2)\right]^2}. \tag{2}$$

This we will assume without proof.* If now $k = 0$, the form (2) is as simple as possible. Assuming for the present that $k$ is not zero, we shall proceed to find another form for the line element. For that purpose, we shall first make the substitutions

$$y_1 = \frac{2u}{t+1}, \quad y_2 = \frac{2v}{t+1}, \quad y_3 = \frac{2w}{t+1}, \tag{3}$$

where $u$, $v$, $w$, and $t$ are connected by the identity

$$1 + k^2(u^2 + v^2 + w^2) = t^2. \tag{4}$$

We have then

$$ds^2 = \frac{du^2 + dv^2 + dw^2 - \frac{1}{k^2}dt^2}{t^2} \tag{5}$$

Finally let us introduce $x$, $y$, and $z$ defined by the relations

$$\left.\begin{aligned} u &= \frac{1}{k}\tan kx, \\ v &= \frac{1}{k}\sec kx \tan ky, \\ w &= \frac{1}{k}\sec kx \sec ky \tan kz \\ t &= \sec kx \sec ky \sec kz. \end{aligned}\right\} \tag{6}$$

It is readily seen that, if $(u_0, v_0, w_0, t_0)$ is any set of values satisfying (4), and if $(b_1, b_2, b_3)$ and $(x_0, y_0, z_0)$ are the corresponding sets of values of $(y_1, y_2, y_3)$ from (3) and $(x, y, z)$ from (6) respectively then the points $(u, v, w, t)$ that lie in the neighborhood of $(u_0, v_0, w_0, t_0)$ and satisfy (4) are transformed by (3) and (6) respectively in a one-to-one manner and continuously on *all* the points $(y_1, y_2, y_3)$ and $(x, y, z)$ respectively that lie in the neighborhood of $(b_1, b_2, b_3)$ and $(x_0, y_0, z_0)$.

The line element now takes on the *normal form*

$$ds^2 = \cos^2 kz(\cos^2 ky\, dx^2 + dy^2) + dz^2. \tag{7}$$

If $k$ is pure imaginary, the functions in (6) and (7) may be written as hyper-

* This form is given by Riemann without proof. It is proved by Dedekind-Weber (*l. c.*) on the assumption that the geodesic surfaces are of constant curvature. Since this assumption has been shown by Schur (*l. c.*) to be warranted, the proof is valid.

A proof has also been given by Lipschitz in connection with extended investigations on the properties of differential expressions; *Crelle*, vols. 70, 71, 72 (1869/70).

bolic functions and the use of imaginaries obviated. We shall avoid duplication of our work, however, by using (6) and (7) as they stand without regard to the reality of $k$, but it may be noticed that $k$ always appears in such combinations that the resulting expressions are real.

The differential equations of the geodesic lines (*A*) become, after a few easy simplifications,

$$\left.\begin{aligned}
\frac{d^2x}{ds^2} &= 2k\tan ky\,\frac{dx}{ds}\,\frac{dy}{ds} + 2k\tan kz\,\frac{dx}{ds}\,\frac{dz}{ds},\\
\frac{d^2y}{ds^2} &= 2k\tan kz\,\frac{dy}{ds}\,\frac{dz}{ds} - k\cos ky\,\sin ky\left(\frac{dx}{ds}\right)^2,\\
\frac{d^2z}{ds^2} &= -k\cos kz\,\sin kz\left[\cos^2 ky\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2\right].
\end{aligned}\right\} \tag{8}$$

*The Coordinates* $(x, y, z)$. The coordinates $(x, y, z)$ introduced analytically by equations (3), (4), (6) are analogous to Cartesian coordinates. For, we can see by inspection that all lines of the three types:

$$\begin{aligned}
&(a) \qquad x = \text{const.},\ y = \text{const.},\ z = s:\\
&(b) \qquad x = \text{const.},\ y = s,\ \ z = 0:\\
&(c) \qquad x = s,\ \ y = 0,\ \ z = 0:
\end{aligned}$$

satisfy equations (8) and are therefore geodesic lines. In particular, the three lines $(x = s,\ y = 0,\ z = 0)$, $(x = 0,\ y = s,\ z = 0)$ and $(x = 0,\ y = 0,\ z = s)$ form a set of three mutually perpendicular geodesic lines intersecting at $O$. We will call them $OX$, $OY$ and $OZ$ respectively. Consider now the surface $z = 0$. We shall have geodesic lines lying on this surface if we place $z = 0$ and take $y$ and $x$ as solutions of the equations

$$\frac{d^2x}{ds^2} = 2k\tan ky\,\frac{dx}{ds}\,\frac{dy}{ds},$$

$$\frac{d^2y}{ds^2} = -k\cos ky\,\sin ky\left(\frac{dx}{ds}\right)^2,$$

which result from placing $z = 0$ in (8). These equations, however, are the same as those which we obtain by applying to the surface $z = 0$, in which the line element is

$$ds^2 = \cos^2 ky\,dx^2 + dy^2$$

the equations of the geodesic lines on any surface, as given in §3, *a*. That is, the shortest line which can be drawn on the surface $z = 0$ connecting any two points is also a geodesic line in space. Hence the surface $z = 0$ is a geodesic surface, any point of which can be taken as the vertex of a generating pencil of geodesic lines. In particular the surface is the geodesic surface determined by $OX$ and $OY$ and may properly be called the surface $XOY$. Similar results may be obtained for the surfaces $y = 0$ and $x = 0$.

Let now $(x_0, y_0, 0)$ be any point on $XOY$. Any geodesic line on $XOY$ through this point has then the direction $(\alpha_1, \beta_1, 0)$, where $\alpha_1^2 \cos^2 k y_0 + \beta_1^2 = 1$. Any line of type $(a)$ passing through $(x_0, y_0, 0)$ has the direction $(0, 0, 1)$. Hence the lines of type $(a)$ are perpendicular to the pencil of geodesic lines on $XOY$ through the point in which $(a)$ meet $XOY$, and are therefore perpendicular to $XOY$. Similarly the lines of type $(b)$ lie in $XOY$ and are perpendicular to $OX$. The only line of type $(c)$ is $OX$. It follows then that if we measure along $OX$ the distance $x_0$, from the point thus reached measure along a geodesic line perpendicular to $OX$ the distance $y_0$, and from the point thus attained measure along a geodesic perpendicular to $XOY$ the distance $z_0$ we reach a point with coordinates $(x_0, y_0, z_0)$. Thus every point $(x_0, y_0, z_0)$ lying in the neighborhood of the point $(0, 0, 0)$ can be reached along one of the paths just described, and along but one such path. As a consequence of the foregoing we have the theorem:

*Through any point P lying in the neighborhood of the point $x = y = z = 0$ one perpendicular to the geodesic surface XOY can be drawn and, if we require the perpendicular to lie wholly in the neighborhood of the above point, but one. From any such point P lying in the geodesic surface XOY one perpendicular to the geodesic line OX can be drawn and, under the same restriction as before, but one.*

To sum up then: *The coordinate system $(x, y, z)$ involves three mutually perpendicular geodesic lines OX, OY, OZ intersecting at O, and determining three geodesic surfaces XOY, XOZ, YOZ. The coordinates of a point P are determined by taking a geodesic line through P perpendicular to XOY meeting it at M, and from M drawing a geodesic line in XOY perpendicular to OX meeting it at N. The lengths of the lines ON, NM, MP are then x, y and z respectively.*

For the sake of symmetry the three geodesic lines and the three surfaces are specified. Practically only one surface $XOY$, one line $OX$, and the point $O$ are needed.

**6. The Equations of the Geodesic Lines and Surfaces.** To obtain the equations of the geodesic lines we need to integrate the differential equations (8), having regard to (7), or its equivalent,

$$1 = \cos^2 ky \cos^2 kz \left(\frac{dx}{ds}\right)^2 + \cos^2 ky \left(\frac{dy}{ds}\right)^2 + \left(\frac{dz}{ds}\right)^2. \tag{9}$$

This may be done by considering the functions derived from (6),

$$\left.\begin{aligned} \xi &= \frac{u}{t} = \frac{1}{k} \sin kx \cos ky \cos kz, \\ \eta &= \frac{v}{t} = \frac{1}{k} \sin ky \cos kz, \\ \zeta &= \frac{w}{t} = \frac{1}{k} \sin kz. \end{aligned}\right\} \tag{10}$$

By linear combinations of (8) and (9), we may obtain the equations:

$$\left.\begin{aligned} \frac{d^2\xi}{ds^2} + k^2\xi &= 0, \\ \frac{d^2\eta}{ds^2} + k^2\eta &= 0, \\ \frac{d^2\zeta}{ds^2} + k^2\zeta &= 0. \end{aligned}\right\} \tag{11}$$

To verify this it is only necessary to multiply equations (8) and (9) by the undetermined factors $\lambda, \mu, \nu, \rho$, to add, and to compare the results with equations (11), written out in terms of $x, y, z$ and their differential coefficients. The values of the multipliers are then easily found. *Hence any solution of equations* (8) *which satisfies* (9) *satisfies also* (11).

Conversely, equations (8) may be expressed as linear combinations of (11) and (9). *Hence any solution of* (11) *which satisfies* (9) *also satisfies* (8).

The general solution of (11) is

$$\left.\begin{aligned} \xi &= \frac{A_1}{k} \sin ks + B_1 \cos ks, \\ \eta &= \frac{A_2}{k} \sin ks + B_2 \cos ks, \\ \zeta &= \frac{A_3}{k} \sin ks + B_3 \cos ks. \end{aligned}\right\} \tag{12}$$

It is necessary to choose the constants so as to satisfy (9). The direct substitution is difficult, but we may proceed indirectly as follows. We notice first that (4) gives us, as a relation which $t$ must satisfy, the following:

$$\xi^2 + \eta^2 + \zeta^2 = \frac{1}{k^2} - \frac{1}{k^2 t^2} \tag{13}$$

Furthermore, (9) is equivalent to (5), and the latter may be written:

$$\left(\frac{d\xi}{ds}\right)^2 + \left(\frac{d\eta}{ds}\right)^2 + \left(\frac{d\zeta}{ds}\right)^2 = 1 - \frac{1}{k^2 t^4}\left(\frac{dt}{ds}\right)^2. \tag{14}$$

Any solution of (11) is also a solution of

$$\xi\frac{d^2\xi}{ds^2} + \eta\frac{d^2\eta}{ds^2} + \zeta\frac{d^2\zeta}{ds^2} + k^2(\xi^2 + \eta^2 + \zeta^2) = 0,$$

which is the same as

$$\tfrac{1}{2}\frac{d^2}{ds^2}(\xi^2 + \eta^2 + \zeta^2) - \left[\left(\frac{d\xi}{ds}\right)^2 + \left(\frac{d\eta}{ds}\right)^2 + \left(\frac{d\zeta}{ds}\right)^2\right] + k^2(\xi^2 + \eta^2 + \zeta^2) = 0.$$

Hence we get from (13) and (14), as a necessary condition for $t$, the equation:

$$\frac{d}{ds}\left(\frac{1}{t}\frac{dt}{ds}\right) = \left(\frac{1}{t}\frac{dt}{ds}\right)^2 + k^2.$$

From this

$$\frac{1}{t}\frac{dt}{ds} = k\tan k(s + a),$$

where $a$ is the constant of integration. No generality is lost by taking $a = 0$, since this amounts to denoting by $s$ the length of the arc measured from a particular point, plus a constant; or simply, as will appear later, to measuring the arc from an arbitrary point. Placing $a = 0$ and integrating the last equation we have

$$t = c\sec ks \tag{15}$$

as a necessary condition for the form of the function $t$.

It follows from the foregoing that, if the equations (12) are to satisfy (9), they must be consistent with (15). Comparison with (14) gives as necessary relations between the coefficients

$$A_1^2 + A_2^2 + A_3^2 = \frac{1}{t^2} + k^2(B_1^2 + B_2^2 + B_3^2) = 1,$$

$$A_1B_1 + A_2B_2 + A_3B_3 = 0.$$

[TO BE CONTINUED.]

# CONTENTS.


| | PAGE |
|---|---|
| Some Applications of the Method of Abridged Notation. By Professor Maxime Bôcher, . . . . . | 45 |
| On the Roots of Functions connected by a Linear Recurrent Relation of the Second Order. By Doctor M. B. Porter, . . . | 55 |
| Space of Constant Curvature. By Professor Frederick S. Woods, | 71 |



ANNALS OF MATHEMATICS.

Published in October, January, April, and July, under the auspices of Harvard University, Cambridge, Mass., U. S. A.

*Cambridge:* Address *The Annals of Mathematics*, 2 University Hall, Cambridge, Mass., U. S. A. Subscription price, $2 a volume (four numbers) in advance. Single numbers, 75c. All drafts and money orders should be made payable to Harvard University.

*London:* Longmans, Green & Co., 39 Paternoster Row. Price, 2 shillings a number.

*Leipzig:* Otto Harrassowitz, Querstrasse 14. Price, 2 marks a number.

Printed by The Salem Press Co., Salem, Mass., U. S. A.